

GenCore version 4.5

Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 7, 2002, 19:49:46 ; Search time 1914.27 Seconds  
 (without alignments)  
 12444.496 Million cell updates/sec

Title: US-09-993-874-1

Perfect score: 1765

Sequence: 1 tcagtgcaagactttaggga.....aaaaaaaaaaaaaaaaaaaaa 1765

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*

2: em\_esthum:\*

3: em\_estin:\*

4: em\_estmu:\*

5: em\_estov:\*

6: em\_estpl:\*

7: em\_estro:\*

8: em\_htc:\*

9: gb\_estl:\*

10: gb\_est2:\*

11: gb\_htc:\*

12: gb\_gss:\*

13: em\_gss\_hum:\*

14: em\_gss\_inv:\*

15: em\_gss\_pln:\*

16: em\_gss\_vrt:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	211	12.0	768	10	BM412281
2	186.8	10.7	721	10	BM413189
3	183	10.4	664	10	C83857
C 4	182.8	10.4	741	9	AU214129
5	180	10.2	882	10	BG45463
C 6	179	10.1	611	10	BJ129072
C 7	179	10.1	667	9	AI486799
C 8	178.2	10.1	604	9	AU219684
C 9	176.2	10.0	638	9	AU214473
10	175.8	10.0	582	9	AW621420
C 11	174.4	9.9	722	10	BJ139943
12	174.2	9.9	606	10	BI923203
13	174.2	9.9	651	10	C90519
14	171.2	9.7	612	10	BG130292
15	170	9.6	722	10	BM413019
16	169.8	9.6	713	10	BM406357
17	169.8	9.6	743	10	BM113456

18	168.6	9.6	569	9	AW218471
19	167.6	9.5	641	10	BG596946
20	167.4	9.5	522	9	AI488821
21	167.2	9.5	625	9	AW776080
22	167.2	9.5	631	10	BE435449
23	166	9.4	570	10	BG791252
24	165.6	9.4	640	10	BM110843
25	165.4	9.4	506	9	AW132800
26	164.8	9.3	629	10	BM143081
27	164.8	9.3	625	9	AI485586
C 28	163.8	9.3	558	9	AW680327
29	162.8	9.2	590	9	AW329421
30	162	9.2	1066	9	BE036528
31	161.2	9.1	520	10	BI269977
32	160.6	9.1	876	10	BG414540
33	160.4	9.1	593	10	BM157962
34	158.2	9.0	656	9	AW691106
C 35	156.6	8.9	694	9	AV860464
36	156.4	8.9	547	9	AV916926
37	156.2	8.8	535	9	AL389862
C 38	156.2	8.8	808	9	AW348381
39	155.6	8.8	644	10	BI309413
C 40	154.8	8.8	772	10	BE820931
41	153.2	8.7	618	9	AW776415
42	152.8	8.7	764	9	AA979869
43	152	8.6	497	9	BE020126
44	151.8	8.6	533	10	BE050505
C 45	151.4	8.6	683	10	BF492687

## ALIGNMENTS

RESULT 1

BM412281

LOCUS

DEFINITION

ACCSSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

Seq primer: T3.

Location/Qualifiers

1. 768

/organism="Lycopersicon esculentum"

/cultivar="TA496"

/db\_xref="taxon:4081"

/clone="cLEG59f7"

/tissue\_type="pericarp"

/dev\_stage="breaker"

/lab\_host="SOLR"

/note="Vector: pBluescriptSKmCUadapt; Site\_1: EcoRI; Site\_2: XhoI; supplier: Boyce Thompson Institute;

BM412281 768 bp mRNA linear EST 22-JAN-2002  
 EST586608 tomato breaker fruit Lycopersicon esculentum cDNA clone  
 CLEG59f7 5' end, mRNA sequence.

BM412281  
 BM412281.1 GI:18263911  
 EST.  
 tomato.  
 Lycopersicon esculentum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Asteridae; euasterids I; Solanales; Solanaceae; Solanum;

1 (bases 1 to 768)  
 Alcaldia, J., Vrebalov, J., White, R., Vision, T., Karamycheva, S.A., Tsai, J., Bougri, O., Kirkness, E., Otterback, T., Van Aken, S., Ronning, C.M., Fraser, C.M., Martin, G.B., Tanksley, S.D. and Giovannoni, J.

Generation of ESTs from tomato fruit tissue, breaker stage (2002)  
 Unpublished (2002)

Contact: CUGI  
 Clemson University Genomics Institute

100 Jordan Hall, Clemson, SC 29634, USA  
 Email: <http://www.genome.clemson.edu/orders/index.html>

This clone is available through the Clemson University Genomics Institute

Seq primer: T3.

Location/Qualifiers

1. 768

/organism="Lycopersicon esculentum"

/cultivar="TA496"

/db\_xref="taxon:4081"

/clone="cLEG59f7"

/tissue\_type="pericarp"

/dev\_stage="breaker"

/lab\_host="SOLR"

/note="Vector: pBluescriptSKmCUadapt; Site\_1: EcoRI; Site\_2: XhoI; supplier: Boyce Thompson Institute;



```

Qy 1381 aaactgcttttacaacatcccaagtattgctgtagctgggttactgaggttcccgagca 1440
Db 428 ACAACTCTTCAATCTCACCTGAAATAGTAGTGTCTGTCTCATACCATATCCTGTGATGA 487

Qy 1441 atttggtgacaattacactgctgctgtgtgtgtgttagaattctggcaagcgcgtga 1500
Db 488 AGAAGCTGGTCAACTGCCATGGCGGTTGTGTGAGAGCTCCCCAAGCACTCTTTGACAA 547

Qy 1501 aaaggaagttcaagattttattgtagcagcaagtcactcccaacaagcatcttcaggcgg 1560
Db 548 AGAACAGTGATGATTTATTTATTTTCAAAACAGGTTGCTCCATATAAGAAGAT---AAGCG 604

Qy 1561 tgcgtattttagacagtattccgaagggccctactgtaaaaaactcatcagaagaagact 1620
Db 605 GGTGGCATTTGTAGTCCATACCAAAAGTCCATCAGGGAAGATATTGAGAAGAGAAT 664

Qy 1621 ccgagaaatattgccca 1638
Db 665 AAAGGATTCATTTACCA 682

RESULT 3
LOCUS C83857 664 bp mRNA linear EST 28-APR-1999
DEFINITION C83857 Dictyostelium discoideum SS (H.Urushihara) Dictyostelium
discoideum cDNA clone SSA391, mRNA sequence.
ACCESSION C83857
VERSION C83857.1 GI:2706789
KEYWORDS EST.
SOURCE Dictyostelium discoideum.
ORGANISM Dictyostelium discoideum.
REFERENCE 1 (bases 1 to 664)
AUTHORS Morio,T., Urushihara,H., Saito,T., Ugawa,Y., Mizuno,H., Yoshida,M.,
Yoshino,R., Mitra,B.N., Pl,M., Sato,T., Takemoto,K., Yasukawa,H.,
Williams,J., Maeda,M., Takeuchi,I., Ochiai,H. and Tanaka,Y.
The Dictyostelium developmental cDNA project: generation and
analysis of expressed sequence tags from the first-finger stage of
development
JOURNAL DNA Res. 5 (6), 335-340 (1998)
MEDLINE 99156227
COMMENT Contact: Hideko Urushihara
Institute of Biological Sciences
University of Tsukuba
3-3-10 Ten-noda1, Tsukuba, Ibaraki 305, Japan
Email: d402huesakura.cc.tsukuba.ac.jp
PROJECT = 'Dictyostelium discoideum cDNA project in Japan'
POLYA-No.

FEATURES
source 1..664
/organism="Dictyostelium discoideum"
/strain="AX4"
/db_xref="taxon:44689"
/clone="SSA391"
/clone_lib="Dictyostelium discoideum SS (H.Urushihara)"
/dev_stage="slug"
BASE COUNT 234 a 105 c 124 g 201 t
ORIGIN

Query Match 10.4%; Score 183; DB 10; Length 664;
Best Local Similarity 55.6%; Pred. No. 2.7e-17;
Matches 351; Conservative 0; Mismatches 280; Indels 0; Gaps 0;

Qy 1027 acaaggatgattgaactgaaactgctgctgctgaatgattaccctcataatgctgt 1086
Db 1 ACAAGTTATGGTGTACCGAAGCTTAGTCCAGCCCTGTTTGTAAATCCAAAGTGGATTAGT 60

Qy 1087 gaaaacaggttcaactggaagacccttgcctacatacattaaagctaaagttttagataaacq 1146
Db 61 TAAATCTGTTCCGCTGTATCCTCTTACCCAATCAATGGCTAAATCATCTCTCCAGA 120

```

```

Qy 1147 tactgggaagcgctaggaccagagaaagagggcaaatatgattcttcaaaagtgaatgat 1206
Db 121 GACTGGTGAGAAATTTAGTGTGAAAGGGTGAAATCTGTATTAAGAGTCCAAATGT 180

Qy 1207 tatgaagaagatatlacaacaatccggaagcaactattgatactattgacaagaatggttg 1266
Db 181 TATGTTGGTTATTATACAAATGAAAGGCAACCAATCAAGTCATAGATAAAGATGGATT 240

Qy 1267 gcttcattctggagatattgatatattacgagaagatggaattcttcttatagtgcagc 1326
Db 241 CTTAAAAAAGTGTGATATTGGTTAGTCGATGAAGATGGTTATTACTTTATCTGTCGATAG 300

Qy 1327 attaaagaacttattaaatacaaggaatatacagtttcgctgctgcaactggaataatct 1386
Db 301 ATCAAAGGAATTTGATCAAAATGTAAGGTTTCCAAAGTACCTCTCGCGAATTTGGAGGCATT 360

Qy 1387 gcttttacaacatcccaagtattgctgtagcggtgttactgagttccggagcgaatttgg 1446
Db 361 ACTATTATCTCATCCAAAAGTTGCAGATGCCTGTGTAGTAGTCTTTCAAAAAGGTGATAT 420

Qy 1447 tggacaattacactgctgctgtgtgtgttagaattctggcaagacgcgtgactgaaaaaga 1506
Db 421 GGGTGAAGTACCAAGAGGTTTCGTTGTTATTAAACAAAATGAATCTCTAACCCGAAAAAGA 480

Qy 1507 agttcaagattttattgagcacaagtcactcccaacaagcatcttcgagcggtgtcgt 1566
Db 481 ACTCTTAGATTGGGCCCCATCCAAAATTTGCAAAATTTATAAACATTTTCAGAGGTGGTATCT 540

Qy 1567 attgtgacagattatccgaaaggccctactggaataactcatcagaagagcgtccgaga 1626
Db 541 TTTCATACCTGCAATTCCTAAATCTGCAACCCGTAACCTATTACGTAAAAATCTTTAAGA 600

Qy 1627 aatattgccagcgagcaccacaaatacaaaa 1657
Db 601 TTTTAATCCTCCAAAATATTAAAAATAAATA 631

RESULT 4
LOCUS AU214129/c 741 bp mRNA linear EST 17-JUL-2001
DEFINITION AU214129 unpublished oligo-capped cDNA library, stage L2
Caenorhabditis elegans cDNA clone yk809f05 3', mRNA sequence.
ACCESSION AU214129
VERSION AU214129.1 GI:14852286
KEYWORDS EST.
SOURCE Caenorhabditis elegans.
ORGANISM Caenorhabditis elegans.
REFERENCE 1 (bases 1 to 741)
AUTHORS Kohara,Y., Shin-i,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.
and Sugano,S.
A complementary view of the C.elegans genome
JOURNAL Unpublished (2001)
COMMENT Contact: Yuji Kohara
Genome Biology Lab.
National Institute of Genetics
Yata 1111, Mishima, Shizuoka 411, Japan
Tel: 81-559-81-6854
Fax: 81-559-81-6855
Email: ykohara@lab.nig.ac.jp.
FEATURES
source 1..741
/organism="Caenorhabditis elegans"
/strain="N2"
/db_xref="taxon:6239"
/clone="yk809f05"
/clone_lib="unpublished oligo-capped cDNA library, stage
L2"
/tissue_type="whole animal"
/dev_stage="L2"
BASE COUNT 216 a 158 c 125 g 241 t 1 others

```



Db 729 GAAATCAAGCATATATTTCGAAACAGGT 57

RESULT 6  
BJ129072/c

LOCUS  
DEFINITION  
Caenorhabditis elegans cdna library, C. elegans L1 stage  
Bj129072  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

FEATURES  
source

BASE COUNT 175 a 131 c 108 g 192 t 5 others  
ORIGIN

Query Match 10.1%; Score 179; DB 10; Length 611;  
Best Local Similarity 60.9%; Pred. No. 1e-16;  
Matches 325; Conservative 0; Mismatches 203; Indels 6; Gaps 2;

QY 1079 aatgctgtgaacagattcaactggaaagaccctgcatacattagaagcttaagt 1138  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
DB 606 AATGATCAACCATATGGTCCGTTGGAAAACCTNGCTCNAATTTGGTTATGAAGATTGG 547  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
QY 1139 gataacgctactggaggcgcttagaccagagagaaagcgaaatatgtttcaaat 1198  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
DB 546 GANCCAGGAACCCACAGACAGAACCGGTTACCACAGGAGAGAATTTTGTCGTGGT 487  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1199 gaattgattgaaggatattacaacaatccggaaagcaactattgatatttgacaaa 1258  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
DB 486 CCAACAATATATGCTGGTATCTCGGAACACAGAGCAAC--AGCCAGCACATGTCATC 430  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1259 gatggttgccttcattcctggagatattgatattacgacgaagattggaattctttata 1318  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
DB 429 GATGGATGGGTTCCACATCGAGACATGTTGTTATCAATGAGGATGGAATTTGTTTTATT 370  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1319 gtgatcatgtgaagaacttatataacaggatcacaggttgccgcttgctgaacctg 1378  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
DB 369 GTTGATGCAGCTTAAGAGCTAATCAAGGTCGAAGGGCTTCAAGTACCACAGCTGAACCT 310  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1379 gaaaaatctgttttaacaacatccaagtatttgtcggtgttactggaggttcggac 1438  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
DB 309 GAAGATCTTCTATCTATCTATCCAAAGATTAGAGATTGTGCGATTATTGGAATCCAGAT 250  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1439 gaatttggtagaacataactcgtcgttctgtgtgtgttagaatctggcaagcgtgact 1498  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
DB 249 GCTFAAGACTGGAGATTAACCTAAGCATTTGTTG---TAAGAGCTGATAACACTCTCACT 193  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1499 gaaaaggaagttcaagatttttttcagcacaaagtctactccccaaaagacatcttcaggc 1558  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
DB 192 GAACAAGAAGTTAAGCATTCTTAAACCAAAGTGTCTCTCTTACAAACACTCGAAGGT 133  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1559 ggtgtcgtattttagacagtagtattccgaaagccctactggaataactcatcaga 1612  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
DB 132 GGTGTCCAATTTATTGAAGAAATTCCAAAGTCCGAGCTGGAAGAGATTTTGAGA 79  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 7  
AI486799  
LOCUS  
DEFINITION  
EST245121 tomato ovary, TAMU Lycopersicon esculentum cDNA clone  
CLEP1D103, mRNA sequence.  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

FEATURES  
source

BASE COUNT 201 a 120 c 155 g 191 t  
ORIGIN

Query Match 10.1%; Score 179; DB 9; Length 667;  
Best Local Similarity 55.8%; Pred. No. 1e-16;  
Matches 372; Conservative 0; Mismatches 280; Indels 15; Gaps 1;

QY 848 cagaattacaaaatcccactatttagtggccctccagtttagtggtttttggtctaaa 907  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
DB 1 CAANAATTACAAGTGACCATAGGGCCATTTGTGCCACCTATTGTTTGGCTATTGCTAAG 60  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 908 agccattagtgcatacaatcatttatcagcttaacggaagtgtctactcggaggagct 967  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
DB 61 AGTCCTATGTTGTTAATTTATGATTATTCATCGTGAAGAACCGTTATGCTGGGGCTGCA 120  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 968 cctttgaaagagatgtcgcagaagcagtagcaagagggttgaaattacciggaatcata 1027  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
DB 121 CCATTAGGAAGAAGACTTGAAGACACTGTTCCAGGCCAATTTCTTAATGCTTAACCTTGGT 180  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1028 caaggataggattaactgaa-----acttgctgcgtgtgaatgattacc 1072  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```
Db 181 CAAGGTTACGGAATGACAGACCGGACCAAGTGTGGCTATGTGCTGGCAATTTGCTAAA 240
Qy 1073 cctcataatgctgtgaacacaggttcaactgtgaagacccttgccatacatataaagctaaa 1132
Db 241 GAACCCTTTGAATAATCTGGACATCTGGACTGTGTGTAGAAATGCTCAGATGAAA 300
Qy 1133 gttttagataacgtactgtgaagcgctgagaccagagagaagcggaatatgtcttt 1192
Db 301 ATTGTGGATCTGATACCTGGCAAAATCTCTCCCTAGGAACCAATCTGGAGAGATTTGTATA 360
Qy 1193 caaagtgaatgattatgaaggatattacaacataccggaacccggaagaactatgatactatt 1252
Db 361 AGAGCGCATCAAAATTTATGAAGGTTTACCTAAATGATGATGACAGGCCACTACGGGAACAATA 420
Qy 1253 gacaaagatggttggctctcatcttgagatatattgatatattacgcgaagatggaaatttc 1312
Db 421 GACAAAGAAAGGTGTTACATACGGCCGATATTGTTATATTGACATCGATGATGAGCTT 480
Qy 1313 ttatagttgatcgattgaagaactattataatacaaggagatatcaggttgcgctgct 1372
Db 481 TTCAATTGTGGATCGTTTAAAGGAATTGATAAAATACAAAGGATTTCAAGTGGCTCCTGCT 540
Qy 1373 gaactggaaaatctgcttttcaacatccaagtattgctgtagtgaggggttacttgagtt 1432
Db 541 GAACTTGAAGCCCTTCTCCTCAATCATCCCAATATTTTCAGATGCTGCTGTTTCAATG 600
Qy 1433 cgggacgaatttggtagcaattacactgctgctgtgtgtgttagaatactcgggaagacg 1492
Db 601 AAGAGCAGCAGCAGCAGGAGAGTTCCTCTGCTGCTTTTGTGTCAGATCAAAATGGCTCCACA 660
Qy 1493 ctgactg 1499
Db 661 ATTACTG 667

RESULT 8
AU219684/c 604 bp mRNA linear EST 17-JUL-2001
LOCUS AU219684
DEFINITION AU219684 unpublished oligo-capped cDNA library, stage L1
Caenorhabditis elegans CDNA clone yk880b12 3', mRNA sequence.
ACCESSION AU219684
VERSION AU219684.1 GI:14857841
KEYWORDS EST.
SOURCE Caenorhabditis elegans.
ORGANISM Caenorhabditis elegans
REFERENCE 1 (bases 1 to 604)
AUTHORS Kohara,Y., Shin-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.
and Sugano,S.
TITLE A complementary view of the C.elegans genome
JOURNAL Unpublished (2001)
COMMENT Contact: Yuji Kohara
Genome Biology Lab.
National Institute of Genetics
Yata 1111, Mishima, Shizuoka 411, Japan
Tel: 81-559-81-6854
Fax: 81-559-81-6855
Email: ykohara@lab.nig.ac.jp.
Location/Qualifiers
1. .604
/organism="Caenorhabditis elegans"
/strain="N2"
/db_xref="taxon:6239"
/clone="yk880b12"
/clone_lib="unpublished oligo-capped cDNA library, stage
L1"
/sex="Hermaphrodite"
/tissue_type="whole animal"
/dev_stage="L1"

BASE COUNT 177 a 129 c 107 g 190 t 1 others
ORIGIN
```







```
Db 136 ACCTTTCATCAGTAGGATAGTCATTTCTGGGGCAGCACCATTAGCAAGAAGCTTGAAG 195
Qy 990 aagcagtagcaaaagggttgaaattcacctggaatcatcaacaggatattgattaaactga-- 1047
Db 196 AGGTAGTTGGAAATTAATTTCTTAATGCTAAACTTGGTCAGGGTTATGGAATGACGGAGG 255
Qy 1048 -----aactgctgcgtgtaagtattaccctctcataatgctctgaaaaaacg 1094
Db 256 CCGGCGACACCTATAACAATGCTTTAGCATTTGCAAAAAGAACCTTTTGAGATAAATCAG 315
Qy 1095 gtcaactggaagacccttgccatcacattaaagctaaagtatttagataacgctactggga 1154
Db 316 GGGCATGTGGCACAGTATATAGGAATGCTGAGATGAAATTTGTGTGATCCAGATCTGGTA 375
Qy 1155 aggcgctaggaccaggagaagagcgaaatagctttcctcaagtgaatgattatgaag 1214
Db 376 TTTCTCTTCTCGAAATAAACCCGGAGAAATTTGCATTAGAGGTGACCAAAATATTGAAAG 435
Qy 1215 gatattacaacaatccggaagcaaatattgatactattgacaaagatggttgcttcatt 1274
Db 436 GTATTTAAATGATCCTGAGCGACCTAAAGAACCAATAGATAAAGGATGGTTACATA 495
Qy 1275 ctggagatattgattaccagcagagatggaaattctttatagttgattgattgaaag 1334
Db 496 GTGGTGATATAGGCTATATTGACGATGATGATGACCTTTTATTGTGGATCGATTAAAGG 555
Qy 1335 aacttattbaatacaaggagattcaggttgcgcctgctgaaactggaa 1381
Db 556 AATTGATAAATAAAGGATTTCAAGTAGCAGCCTGCTGAGCTTCAA 602

RESULT 13
C90519 651 bp mRNA linear EST 20-APR-1998
LOCUS C90519 Dictyostelium discoideum SS (H.Urushihara) Dictyostelium
DEFINITION discoideum cDNA clone SSI661, mRNA sequence.
ACCESSION C90519
VERSION C90519.1 GI:3060139
KEYWORDS EST.
SOURCE Dictyostelium discoideum.
ORGANISM Dictyostelium discoideum
REFERENCE Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
AUTHORS Morio,T., Urushihara,H., Saito,T., Ugawa,Y., Mizuno,H., Yoshida,M.,
Yoshino,R., Mitra,B.N., Pl,M., Sato,T., Takemoto,K., Yasukawa,H.,
Williams,J., Mada,M., Takeuchi,I., Ochiai,H. and Tanaka,Y.
TITLE Developmental cDNA in Dictyostelium discoideum
JOURNAL Unpublished (1998)
COMMENT Contact: Hideko Urushihara
Institute of Biological Sciences
University of Tsukuba
3-3-10 Ten-nodai, Tsukuba, Ibaraki 305, Japan
Email: d402huesakura.cc.tsukuba.ac.jp.
FEATURES
source
1. .651
/organism="Dictyostelium discoideum"
/strain="AX4"
/db_xref="taxon:44689"
/clone="SSI661"
/clone_lib="Dictyostelium discoideum SS (H.Urushihara)"
/dev_stage="slug"
BASE COUNT 251 a 83 c 110 g 207 t
ORIGIN

Query Match 9.9%; Score 174.2; DB 10; Length 651;
Best Local Similarity 55.6%; Pred. No. 4.9e-16;
Matches 355; Conservative 0; Mismatches 283; Indels 1; Gaps 1;
Qy 1127 gctaaagttagataacgctactggagcgctagaccaggaagagcgaaata 1186
Db 1 GCTAAATCATCTCTCTAGAGACTGGTGAGAAATTTAGGTATGGGTGAAGAGGCTGAATC 60
```

```
Qy 1187 tgccttcaaaagtgaatgattatgaagagatattacaacaatcccggaagcaactattgat 1246
Db 61 TGTATTAAAGGTCCAAATGTTATGTTGGTTATTATAACAATGAA-AGCAACCAATGAA 119
Qy 1247 actattcaaaagatggttgccttcattctgagatattgatatattacgacaaagatgga 1306
Db 120 GTCATAGATAAAGATGGATTCCTTAAAACTGGTGATATTTGGTTATGTGATGAAGATGT 179
Qy 1307 aattctttatagttgacgattgaagaactattataacaagggatcatcaggttgcg 1366
Db 180 TATTTCTTTATCATCGATAGATCAAAAGGAATGTCAAAATGTAAGGTTTCCAAGTACCA 239
Qy 1367 cctgctgaactggaaaaactcgtcttttaacaacatccaaagtattgctgagcggtgttact 1426
Db 240 CCGTGTGAATTAGAAGCATTACTATTATCTCATCCAAAAGTTGCAGATGCCTGTGTAGTA 299
Qy 1427 gtagtctccggagcaatttggcgacaattaccctcgtcgtctgtgtgtgttagaactggc 1486
Db 300 GGTCTTTCAAAAGGTGATATGGGTGAAGTACCAAGAGGTTTCGTTGTTATTTAAACAAAT 359
Qy 1487 aagacgctgactgaaaaaggaagtccaagattttattgcagcacaagtcactccaacaag 1546
Db 360 GAATCTCTTACTGAAAAGAACTCTTTGGATTGGGCTCACCCAAAATTTGCAAAATTATAAA 419
Qy 1547 catcttcagggcggtgctgatttttagacagatattccgaaaggccctactactggaatactc 1606
Db 420 CATTTTCAGAGGTGGTATCTTTTTCATCTGCAATTCCTAAATCTGCAACTGGTAAACTA 479
Qy 1607 atcagaagaggagctccgagaaatatttggccagcagcaccacaaatacaaatataagtc 1666
Db 480 TTACGTAAAAATCTTAAAGATTTTAAATCTTAAATTTTAAATTTTAAAAAAGGTGT 539
Qy 1667 caatgattgctttagttcttaaaatgtataaaacaaagtttttagaacctaatcatctcat 1726
Db 540 TAAATTTCCACTTTCACCAACAAGATTTTCAAAATTAATTTATTATATATAATGTAGGAAA 599
Qy 1727 tcaatatacaacaaaaaiaaaaaaiaaaaaaiaaaaaaiaaaaaa 1765
Db 600 CCAAAATGCCAACAAATAAAATTTTAAAGATATATATATAA 638

RESULT 14
BG130292
LOCUS BG130292 612 bp mRNA linear EST 31-JAN-2001
DEFINITION EST475938 tomato shoot/meristem Lycopersicon esculentum cDNA clone
ctOF29J22 5' sequence, mRNA sequence.
ACCESSION BG130292
VERSION BG130292.1 GI:12630480
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AUTHORS Hansen,C., Ronning,C. and Tanksley,S.
TITLE Generation of ESTs from tomato shoot/meristem tissue
JOURNAL Unpublished (2001)
COMMENT Contact: CUGI
Clemson University Genomics Institute
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html.
location/Qualifiers
1. .612
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="ctOF29J22"
/clone_lib="tomato shoot/meristem"
FEATURES
source
```



Search completed: September 7, 2002, 21:31:54  
Job time: 6128 sec

---

***This Page Blank (uspto)***

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 7, 2002, 20:03:21 ; Search time 77.82 Seconds  
(without alignments)  
5571.102 Million cell updates/sec

Title: US-09-993-874-1  
Perfect score: 1765  
Sequence: 1 tcagtcgaagacttttagga.....aaaaaaaaaaaaaaaaaaaaa 1765

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:\*

- 1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq:\*
- 2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq:\*
- 3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq:\*
- 4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq:\*
- 5: /cgn2\_6/ptodata/2/ina/PCTUS\_COMB.seq:\*
- 6: /cgn2\_6/ptodata/2/ina/backfiles.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	582	33.0	1659	1 US-08-231-729B-1	Sequence 1, Appli
2	581.4	32.9	1659	1 US-08-231-729B-2	Sequence 2, Appli
3	580.4	32.9	1725	4 US-09-380-061B-19	Sequence 19, Appl
4	584.4	31.4	5620	2 US-08-793-170-21	Sequence 21, Appl
5	584.4	31.4	5620	3 US-08-892-873-21	Sequence 21, Appl
6	584.4	31.4	5620	4 US-09-334-765A-21	Sequence 21, Appl
7	584.4	31.4	5620	4 US-09-356-575E-21	Sequence 21, Appl
8	584.4	31.4	5620	4 US-09-333-820-21	Sequence 21, Appl
9	582.8	31.3	1811	3 US-08-867-352-22	Sequence 22, Appl
10	582.8	31.3	5789	3 US-08-862-431-32	Sequence 32, Appl
11	582.8	31.3	5791	3 US-08-862-431-31	Sequence 31, Appl
12	582.8	31.3	5793	3 US-08-862-431-29	Sequence 29, Appl
13	582.8	31.3	5793	3 US-08-862-431-30	Sequence 30, Appl
14	582.8	31.3	5818	2 US-08-536-559A-3	Sequence 3, Appli
15	582.8	31.3	5819	2 US-08-536-559A-2	Sequence 2, Appli
16	582.8	31.3	5819	3 US-08-862-431-27	Sequence 27, Appl
17	582.8	31.3	5819	3 US-08-862-431-28	Sequence 28, Appl
18	582.8	31.3	5938	2 US-08-536-559A-4	Sequence 4, Appli
19	582.8	31.3	6092	2 US-08-536-559A-1	Sequence 1, Appli
20	582.8	31.3	6092	3 US-08-862-431-26	Sequence 26, Appl
21	582.8	31.3	11616	1 US-08-186-259-2	Sequence 2, Appli
22	580.6	31.2	5427	3 US-09-282-996-2	Sequence 2, Appli
23	580.2	31.2	1722	4 US-08-875-277A-1	Sequence 1, Appli
24	550.2	31.2	1722	3 US-08-718-425-1	Sequence 1, Appli
25	549.8	31.2	1650	1 US-08-354-240A-1	Sequence 1, Appli
26	549.8	31.2	1722	4 US-09-380-061B-1	Sequence 1, Appli
27	549.8	31.2	2445	1 US-08-122-520C-8	Sequence 8, Appli

Sequence 1, Appli  
Sequence 18, Appli  
Sequence 18, Appli  
Sequence 12, Appli  
Sequence 12, Appli  
Sequence 17, Appli  
Sequence 17, Appli  
Sequence 17, Appli  
Sequence 17, Appli  
Sequence 17, Appli  
Sequence 15, Appli  
Sequence 15, Appli  
Sequence 13, Appli  
Sequence 8, Appli  
Sequence 8, Appli  
Sequence 5, Appli  
Sequence 5, Appli  
Sequence 8, Appli  
Sequence 7, Appli  
Sequence 7, Appli  
Sequence 6, Appli

ALIGNMENTS

RESULT 1  
US-08-231-729B-1  
; Sequence 1, Application US/08231729B  
; Patent No. 5618722  
; GENERAL INFORMATION:  
; APPLICANT: ZENNO, Shubei  
; APPLICANT: SHIRASHI, Shinji  
; APPLICANT: INOUE, Satoshi  
; APPLICANT: SAIGO, Kaoru  
; TITLE OF INVENTION: FIREFLY LUCIFERASE GENE  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LEYDIG, VOIT & MAYER  
; STREET: 700 Thirteenth Street, N.W., Suite 300  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/231.729B  
; FILING DATE: 20-APR-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 119050/1993  
; FILING DATE: 21-APR-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Rose, Herbert C.  
; REGISTRATION NUMBER: 29846  
; REFERENCE/DOCKET NUMBER: 60130/No. 5618722aka  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-737-6770  
; TELEFAX: 202-737-6776  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1659 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1..1659  
; OTHER INFORMATION: "Xaa" at codon 409 is either Glu or Asp;  
; OTHER INFORMATION: "Xaa" at codons 28, 32, 112, 130, 142, 190, 212, 217, 222,  
; OTHER INFORMATION: 329, 336, 386, 436, 512, and 532 is either Arg, Ser or Gly



Db	666	NGAYCCNGNTTYTGGNAAACARATHATCCNGAYACNACNATHYNTGYCGNGTNCNCNTT	725
Qy	745	ccatcatgcttggaaatgcttactacattacttactttagtaggacttaagcttgc	804
Db	726	YCAYCAYGCTTYTGGNACNTTYACNAAAYTNGNTAYTYTNATHYTCGNTTYCAYGTNGT	785
Qy	805	aatgttgaagaataattgagggcgactttcttataaaaccatacagaataacataaaatccc	864
Db	786	NYTNATGTAYMGNTTYAAYGACAYVTNTTYTNCARACNTYNTNCARGAYTAAARTGYCA	845
Qy	865	cactattgtagtgccctccagttatggttttggcttaaaagccattagtcgatca	924
Db	846	RTCCNGCNYTYNTGTNCCNACNGTYNTGNTTYTTCNAAAYCCNYTNGTNGAYAA	905
Qy	925	atcagattatcgagcttaacggaagtgctactggagagctctttaggaaaagatgt	984
Db	906	RTAYGATYNTCNAAYTNCAGATHGCNTCNGNGNGCNCNTYNTNCNARGARAT	965
Qy	985	cgcagaagcagtagcaaaagaggttgaataacctggaatcatacaagatagattaac	1044
Db	966	HTCNGARATHGCNCNAARMGNTTYAARYTNCNGNATHMGCNARGGNTAYGCNTNAC	1025
Qy	1045	tgaacctgctgcgctgttaattgattaccctccatgaatgctgtgaacaacaggttcaactgg	1104
Db	1026	NGARACNACNTGYCGNATHGTHNATHACNGCNGARGNGARTTYAARYTNGCNGCNGTNGG	1085
Qy	1105	aagacacctgccaataattaaagctaaagtttagataaacgctactgggaagcgctagg	1164
Db	1086	NAARGTNGTNCNCNTTYATCYNTNNAARGTNTNGAYTYTNAAYACNGNGNAAARYTNGG	1145
Qy	1165	accagagaaagagcgaaatctgctttcaaaagtgaatgattgaagagatatataca	1224
Db	1146	NCCNAAYGABMGNGNGARATHGTGTYTYAARGNCCNATGATHATGAARGCNTAYATHAA	1205
Qy	1225	caatccgggaagcaactatgatactatttgcaaaagatggttggcttcaatctgagatata	1284
Db	1206	YAAAYCCNGARGCNACNCGCGARYTNATHGAYGARGARGNTGGATHCAAYTCNGNGAYAT	1265
Qy	1285	tgatattacgcgaagatggaaatctcttatagttgatcgatgaaagaactattaa	1344
Db	1266	HGGNTATYTGAYGARGAYGGNCAYGTNTAYATHGTNCAWMGNTYNTNATCNYTNATHAA	1325
Qy	1345	atacaagggatatacaggttcgcgctgctgaactggaaatctgctttacaacatccaag	1404
Db	1326	RTAYAARGGNTAYCARGTNCNCNCGNARYTNGARGCNYTNTYNTNCARCAAYCNCNTT	1385
Qy	1405	tattgctgatacggggttactggagttccggacgaatttggcgacaattacctgctgc	1464
Db	1386	YATHGARGAYCGNGNGTNGCNGGNTCCNGAYGARGTNGCNGNGAYTYTNCNGGNGC	1445
Qy	1465	ttgtgtgttagaatctggcaagaacgctgactgaaagaagttcaagattttattgc	1524
Db	1446	NGTNGTNGTYTNAAARGNGNARTCNATHACNGARAAARGARATHCARGATAYGTNGC	1505
Qy	1525	agcaagaagtcactccaacaaagcatcttcgagcggtgtcgattgtgacagattccc	1584
Db	1506	NGGNCARGTNACNTCTCNAAARAARYTNGNGGNGGNGTNGARTTYGTNAARGAGTNC	1565
Qy	1585	gaaaggccctatggaaaactcatcagaagagctccgagaaatatttgcgcag	1639
Db	1566	NAARGGNTTYACNGGNAARATHGAYACNMGNAARATHAARGARATHYTNATHAAR	1620
RESULT 3			
US-09-380-061B-19			
; Sequence 19, Application US/09380061B			
; Patent No. 6265177			
; GENERAL INFORMATION:			
; APPLICANT: SQUIRELL, DAVID JAMES			
; WHITE, PETER JOHN			
; LOWE, CHRISTOPHER ROBIN			
; MURRAY, JAMES AUGUSTUS HENRY			

;; TITLE OF INVENTION: ENZYME ASSAY FOR MUTANT FIREFLY LUCIFERASE

;; NUMBER OF SEQUENCES: 21

;; CORRESPONDENCE ADDRESS:

;; ADDRESSEE: NIXON & VANDERHYE P.C.

;; STREET: 1100 NORTH GLEBE ROAD

;; CITY: ARLINGTON

;; STATE: VIRGINIA

;; COUNTRY: U.S.A.

;; ZIP: 22201-4714

;; COMPUTER READABLE FORM:

;; MEDIUM TYPE: Floppy disk

;; COMPUTER: IBM PC compatible

;; OPERATING SYSTEM: PC-DOS/MS-DOS

;; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

;; CURRENT APPLICATION DATA:

;; APPLICATION NUMBER: US/09/380.061B

;; FILING DATE: 25-Aug-1999

;; CLASSIFICATION: <Unknown>

;; PRIOR APPLICATION DATA:

;; APPLICATION NUMBER: PCT/GB98/01026

;; FILING DATE: 7-APR-1998

;; APPLICATION NUMBER: GB 9707468.8

;; FILING DATE: 11-APR-1997

;; ATTORNEY/AGENT INFORMATION:

;; NAME: SADOFF, B. J.

;; REGISTRATION NUMBER: 36,663

;; REFERENCE/DOCKET NUMBER: 124-725

;; TELECOMMUNICATION INFORMATION:

;; TELEPHONE: (703)816-4000

;; TELEFAX: (703)816-4100

;; INFORMATION FOR SEQ ID NO: 19:

;; SEQUENCE CHARACTERISTICS:

;; LENGTH: 1725 base pairs

;; TYPE: nucleic acid

;; STRANDEDNESS: both

;; TOPOLOGY: linear

;; MOLECULE TYPE: cDNA

;; FEATURE:

;; NAME/KEY: CDS

;; LOCATION: 32..1675

;; SEQUENCE DESCRIPTION: SEQ ID NO: 19:

us-09-380-061B-19

Query Match 32.9%; Score 580.4; DB 4; Length 1725;

Best Local Similarity 60.08; Pred. No. 8.2e-113;

Matches 1006; Conservative 0; Mismatches 661; Indels 9; Gaps 2;

Qy 29 gaagaagaaaacattagcagtgagagcgtcctcgtgatatagtcctatcctcggtcgcgca 88

Db 38 GATGCAAAAATATTATGCACGGTCCAGCGCATTTCTATCTCTTTGGAGGATGGAAC TGCT 97

Qy 89 ggacaacaattatccaatcattgtataaaatttgcatacttttctcctgaagcaata----- 142

Db 98 GGAGAACAAATTCGCAAAAGCAATGAAGAGGTATGCACAGGTTCGAGGACAAATTCGCTTTT 157

Qy 143 atcagatgctacacaaatgaagtatacatcgtcctcaaatatttgaaccagctgcgcg 202

Db 158 ACTGATGCCCATGACAGAGTAAATATTACATATTCCGAATATTTTGAATGGCTTGCCGG 217

Qy 203 ttactgtgttagtagaacaatatggcttgaaatgaacaacaattgttgggtgtatgcagt 262

Db 218 TTAGCCGAAACTATGAGAGGTACGGACTTGTGTTTACAACACCATTTGCTGTTGTAGC 277

Qy 263 gaaacaacaataaaacttttttaactcctgctccttgccttatacttaggaataaccagta 322

Db 278 GAAATTCCTCTAGTTTTTTATGCTCTATGCGGGGCTCTATTTATTGGAGTTGGAGTT 337

Qy 323 gcaacatacaatgatgatcacagatggagagtttaactggtcatttgaatatatacaaaa 382

Db 338 GCATCAACAAATGATATTACAAAGCTGAATATACACAGTTTGTCCATATACACA 397

Qy 383 ccaactatcatgttttagttcaagaagaagcactcccgcttattcttgagagtacagcaaat 442

Db 398 CCTACAATAGTATCTCTGTTCAAAAAGAGCGCTGCAAAAAATCCTAGGGGTACAAAAAGAA 457

Qy 443 ctaagtttccataaaaaagtcgttagttatcgatagcagatagcagacataataatggcgttga 502

Db 458 TTACCTATAATTCAGAAAAATTTGTTATTCGGAATTCGAGAGGATTTATATGGGAAACAA 517

Qy 503 tgcgtatctactccttctgcacttatactagcaccaccccttctgcatctgcatcttctaca 562

Db 518 TCTATGTACTCGTTTCATTGAATCTCATTTACCTCGACGTTTAAATGAATATGATTACATA 577

Qy 563 ccaaaagatttgcaccccttgaaaaaatgcatttaattatgtcatcatctcctggacaact 622

Db 578 CCGGATTCATTTCACCGGAAACAGCAACAGCACTTATAATGAATTCATCGGATCTACT 637

Qy 623 ggattgcctaagggtgtagttactagccatagaagtcataaagtcataaagattcgttcatagc 682

Db 638 GGATTGCCCAAGGGAGTTGAGCTTACTCACCANAATGTGTGTAGATTTCCTCAGTCG 697

Qy 683 aggatcccatattatggcaactcgtacggttcccaacaacatacaattcttctccttagtaccg 742

Db 698 AGAGATCCTGTGTTTGGTAATCAAAATTAATCCCGATACTGCGATTTTAACAGTTTATACCA 757

Qy 743 ttccatcatgctcttggaaatgttttaactacattatcttactcttctgtagtgacttaaggtt 802

Db 758 TTTTCATCATGTTTGGAAATGTTTACAACACTAGGATATTTAAGCTGTGGATTTTCGTATT 817

Qy 803 gtaatgttgaagaataatttggaggcgcacttttcttaaaaccatacacagaattacaaaatc 862

Db 818 GTGCTTATGTATAGATTTGAAGAGGAATTAATTTTACGATCATTCAAGATTTATAAAAT 877

Qy 863 cccactatgttagtggccctccagttatgggttttggctaaagccatattagtcgat 922

Db 878 CAAAGTCGGTTCGTGCTACCTACTCTATTTTCATCTTTGCCAANAAGCACCTTACTCGAT 937

Qy 923 caatacagatttctcagcttaaacggaagttgctactgagagagctcctttaggaaaaagat 982

Db 938 AATACGATTTATCCAACTTACATGAAATTCGTCGTGAGCTCCCTCGCGCAAGAA 997

Qy 983 gtccagaagcagtagcaaaagaggttgaataccttgcctggaatcacacagaattatggatta 1042

Db 998 GTTGGAAGCTGTAGCAAAACGTTTAAAGCTGCCGGAATACGACAAAGGTATGGGACTT 1057

Qy 1043 actgaactgtcgtcgtgtaattgattaccctctcataatgctgtgaaacaggttcaact 1102

Db 1058 ACTGAACACTACCTCAGCTATTATTAATTAACACCAAGGGGATGATAACACAGGAGCATGT 1117

Qy 1103 ggaagaccccttgccatcacattaaagctaaagtttttagtaaacgctactggaagagcgcta 1162

Db 1118 GGTAAAGTTGTCCATCTCTTCTGCCAAAATTCCTGATCTGGATACGGGTAAACCTTTG 1177

Qy 1163 ggaccagagagaaagcgcaaatatgctttcnaaagtgaaatgattatgaaaggatattac 1222

Db 1178 GGTGTTAATACAGAGGGGGGAATTAATGTGTAAGGCGCCCAATGATAAATGAAGGGTTACGTA 1237

Qy 1223 aacaatccgaagcaactattgatactattgacaaagatggttggcttcattctcggagat 1282

Db 1238 AACAAACCCAGAGCAACAAAGTGCATTGATACACAAAGATGTTGTTACACTCTCTGGTGAC 1297

Qy 1283 attggattattacgacgaagtggaaatttctttagttgatacgattgaaagaacttattt 1342

Db 1298 ATAGCTTACTACGACAAGATGGTCACCTTCTTCATAGTAGATCGTTTGAATCGTTAAT 1357

Qy 1343 aaatcacaaaggatatacaggttgccctcgtggaactggaaaaatcgttcttacaacatcca 1402

Db 1358 AAATACAAAGGTTATCAGGTACCCCTGCCGAATTTAGAAATCGATATTGCTGCAACATCCC 1417

Qy 1403 agtattgctgatcggtgttactcgtgaggttcggagcgaatttgggtggacaattaccgct 1462

Db 1418 TTCATATTTGATGACAGGTGTTCAGGAATTTCCCGACCCAGATGCCGGTGAACCTTCTTCCA 1477

Qy 1463 gcttgtgtgtgttagaattcgtgcaagacgctgactgaaaggaaggttcgaagatttttatt 1522



Db 1478 GCCGTTCTGCTTAGAGGAGGCAAAACGATGACTGNACAAAGTGGATGATTAATGTT 1537  
Qy 1523 gcagcacaagtcactccaacaaagcatcttcgagcggtgctgtattttagtagacaglatt 1582  
Db 1538 GCGGACAAAGTAACCTTCTTAAGCGTTTACGTGGAGGAGTTAAGTTTGTGGACGAAGTA 1597  
Qy 1583 ccgaaagccctactggaaactcatcagaaagagagctccgagaaatatttggccagcga 1642  
Db 1598 CCTAAGGCTCAACTGGAAAGATTGATGGAAAGAAATTCAGGGAGATGCTT--ATGATG 1654  
Qy 1643 gcacccaaataataataatgctcaatgctatttcttagtctcaaaatgtatata 1698  
Db 1655 GGAANAATCCAAATGTAATTCCTTCGGTTTACTATATATCTAACGAAATTTCTA 1710

## RESULT 4

US-08-793-170-21  
; Sequence 21, Application US/08793170  
; Patent No. 5994128  
; GENERAL INFORMATION:  
; APPLICANT: FALLAUX et al.  
; TITLE OF INVENTION: PACKAGING SYSTEMS FOR HUMAN RECOMBINANT  
; TITLE OF INVENTION: ADENOVIRUS TO BE USED IN GENE THERAPY  
; NUMBER OF SEQUENCES: 22  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: RAE-VENTER LAW GROUP, P.C.  
; STREET: 260 SHERIDAN AVENUE, PO BOX 60039  
; CITY: PALO ALTO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94306  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/793.170  
; FILING DATE: 25-MAR-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: WO 97/00326  
; FILING DATE: 14-JUN-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: EP 95201728.3  
; FILING DATE: 26-JUN-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: EP 95201611.1  
; FILING DATE: 15-JUN-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: RAE-VENTER, BARBARA  
; REGISTRATION NUMBER: 32,750  
; REFERENCE/DOCKET NUMBER: INGE.002.000S  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (650)328-4400  
; TELEFAX: (650)328-3377  
; TELEX: N/A  
; INFORMATION FOR SEQ ID NO: 21:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 5620 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: other nucleic acid  
; DESCRIPTION: /desc = "plasmid"  
US-08-793-170-21

Query Match 31.4%; Score 554.4; DB 2; Length 5620;  
Best Local Similarity 59.2%; Pred. No. 2.7e-107;  
Matches 967; Conservative 0; Mismatches 661; Indels 6; Gaps 1;

Qy 2 cagtgcagacttttaggatcaaaatggaaagaaacattagggcattagggagcgtcct 61

Db 1265 CATTCGGTACTGTGTGTAATAATGGAAGACGCCAAAAACATAAAGAAAGGCCCGCGCCA 1324  
Qy 62 cgtgatataagccactcctgctcgagcagcaacaattatataccaatcattgtataaaatt 121  
Db 1325 TTCTATCTCTAGAGGATGGAACCGCTGGAGAGCAACTGCATAGGCTATGAAGAAATAC 1384  
Qy 122 gcatttttctcctgaagcaa-----taatcgatgctcgcatacaaatgaagtaatatcat 175  
Db 1385 GCCCTGGTTCCTGCAACAATTTGCTTTACAGATGCACATATCGAGGTGAACATCACGTAC 1444  
Qy 176 gctcaaatatttgaaccagcgtcgcttagcttagtagatagacaacaatatggcttgaat 235  
Db 1445 GCGGAATACCTTCGAATAATGCTCGGTTCGAGAGCTATGAACACGATATGGGTGTAAT 1504  
Qy 236 gaaacacaatgtgtgggtgtatgcagtgaaacaaataaaaacttttttaactcctgtcctt 295  
Db 1505 ACAATACACAGATCGTCGATGAGTGAATAACTCTCTCAATCTTATCCGGGTGTTG 1564  
Qy 296 gctgctttatacttaggaatcccgtagcagcaacatacaaatgatatgacacagatggagag 355  
Db 1565 GCGCGTTATTTATCGGAGTTGCAGTTGCGCCGCGGAACGACATTTATATGAACGTGAA 1624  
Qy 356 ttaactggtcatttgaatatatacaaaacaaactatcatgttttagtttcaaaagaaagcactc 415  
Db 1625 TTGCTCAACAGTATGAACATTTCCGAGCTACCGTAGTGTGTTGTTTCCAAAAGGGTGTG 1684  
Qy 416 ccgcttattctgaggtacagcgaacaaactcgaagtttcttaaaaaagtcgtagttatcgat 475  
Db 1685 CAAAAAATTTTGAACGTGCAAAAAAATTTACCAATTAATCCAGAAAAATTTATATCATGAT 1744  
Qy 476 agcatgtacgacataatggtggtgaaatgcgtatctcctttgtgtgacgttatactgac 535  
Db 1745 TCTAAAACGGATTACCGGGATTTTCAGTCGATGTACACGTTTCGTGCACATCTCATCTAC 1804  
Qy 536 cacacctttgatccattgtcattacacaaagattttgatcccttgccttgaaacaaatcgca 595  
Db 1805 CCCGGTTTTAATGAATACGATTTTGTACCAGATGCTTGTGATCGTGACAAACAAATTCGA 1864  
Qy 596 ttaattatgctcatcctggaacactggtgctgcctgaaggtgtagtagtagcagccataga 655  
Db 1865 CTGATAATGAATCTCTCGATCTACTCGGTTACTTAAGGGTGTGGCCCTTCCGCATAGA 1924  
Qy 656 agtctaactataagattcgttcatagcaggagatccactttatgacactcgtacaggttcca 715  
Db 1925 ACTGCTGCGTCAGATCTCGCATGCCAGAGATCTATTTTGGCAATCAATCATCCG 1984  
Qy 716 caaacatcaattcttctcttagtaccgttccatcgcttgcattggaattttactacatta 775  
Db 1985 GATCTGCGATTTTAAGTGTGTTGCCATTCACGTTTGGAAATGTTTACTACATCTC 2044  
Qy 776 tcttactttgttaggacttaaggttgaatgttgaagaaatttgaaggcgccacttttc 835  
Db 2045 GGATATTTGATATGTGGATTTTCGAGTCGCTTAATGTATAGATTTGAAGAGAGCTGTTT 2104  
Qy 836 ttaaaacaccatacagaattacaacaaatcccccactatgttagtggccccctccagttatggt 895  
Db 2105 TTACGATCCCTTCAGGATTACAAAATTTCAAAGTGGTGTGCTAGTACCAACCGCTATTTTCA 2164  
Qy 896 tttttgctaaaagcccatctagtcgatcaatcacgatttctcagcttaacggaagtgtct 955  
Db 2165 TTCTTCGCAAAAGCACCTCTGATTGCAAAATACGATTTATCTAATTTACAGAAATTTGCT 2224  
Qy 956 actggaggagctccttttaggaaaaagatgcgcagagcagtagcagaaggttgaatata 1015  
Db 2225 TCTGGGGCGCACCTCTTTTCGAAAGAACTCGGGGAAGCGGTTGCAAAACGCTTCCATCTT 2284  
Qy 1016 cctggaatcatacaaggatattggaataactgaaacttgcgtgcgtgtaattgattaccct 1075  
Db 2285 CCAGGGATACGCAAGGATATGGGCTCAGTACAGATACATCAGCTATTTCTGATTACACCC 2344  
Qy 1076 cataatgctgtgaaacacaggttcaactggaagaccttgcacatacatataaagattt 1135

Db 2345 GAGGGGATGATAAACCAGGGCGCGTGGTAAAGTGTGTCCATTTTTTGAAGCGAAGGTT 2404  
Qy 1136 ttatataacgctacttggaagcgctagaccaggaagaagaagcggaataatgctttcaa 1195  
Db 2405 GTGGATCTGGATACCGGAAACGCTGGCGTTAATCAGACAGCGGAATTATGTCTCAGA 2464  
Qy 1196 agtgaatgattatgaaggaattatacaacaatccggaagaacaattattgatactattgac 1255  
Db 2465 GGACCTATGATTATGTCGGGTTATGTAACAATCCGGAAGCGACCAACGCCCTTGATTGAC 2524  
Qy 1256 aaagatggttggtcttcattctggagatatgtgatatcagcagaagatggaaaattctttt 1315  
Db 2525 AAGGATGATGGCTACATCTGGAGACATAGCTTACTGGGACGAAGAACACACTTCTTC 2584  
Qy 1316 atagtgtatgatgaagaacttatataacaaaggaatcaggttgccctgctgaa 1375  
Db 2585 ATAGTTGACCCCTTGAAGTCTTAATTAATACAAAGGATATCAGGTGGCCCCCGCTGAA 2644  
Qy 1376 ctgaaaaatctgctttttacaacatccaagtattgctgatcggtgtgttactggagttccg 1435  
Db 2645 TTGGAATCGATATTGTTACAACACCCCAACATCTTCGACGCGGGCTGGCAGGTCTTCCC 2704  
Qy 1436 gacgaatttgtagacaattaccctgctgctgtgtgtgttagaatacttggcaagcgtg 1495  
Db 2705 GACGATGACGCGCGGTGAACCTTCCGCGCGCGCTGTGTGTTTGGAGCACGGAAGACGATG 2764  
Qy 1496 actgaagaaggaattcaagattttatgacgacacaagtcactccaacaagcatcttoga 1555  
Db 2765 ACGMAAAGAGATCGTGGATACGTCGCCAGTCAAGTAACACCGCGGAANAAGTTGCGC 2824  
Qy 1556 ggcgggtgctgatttgtagacagtattccgaaaggccctcattggaactcatcagaag 1615  
Db 2825 GGAGGAGTGTGTTGTGGAGCAAGTACCGAAAGTCTTTACCGGAAACTCGACGCAAGA 2884  
Qy 1616 gagctccagagaat 1629  
Db 2885 AAAATCAGAGAGAT 2898

## RESULT 5

US-08-892-873-21  
; Sequence 21, Application US/08892873  
; Patent No. 6033908  
; GENERAL INFORMATION:  
; APPLICANT: FALLAUX et al.  
; TITLE OF INVENTION: PACKAGING SYSTEMS FOR HUMAN RECOMBINANT  
; TITLE OF INVENTION: ADENOVIRUS TO BE USED IN GENE THERAPY  
; NUMBER OF SEQUENCES: 22  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: RAE-VENTER LAW GROUP, P.C.  
; STREET: 260 SHERIDAN AVENUE, PO BOX 60039  
; CITY: PALO ALTO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94306  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/892.873  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/793.170  
; FILING DATE: 25-MAR-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: WO 97/00326  
; FILING DATE: 14-JUN-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: EP 95201728.3  
; FILING DATE: 26-JUN-1995

; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: EP 95201611.1  
; FILING DATE: 15-JUN-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: RAE-VENTER, BARBARA  
; REGISTRATION NUMBER: 32,750  
; REFERENCE/DOCKET NUMBER: INGE.002.000US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (650)328-4400  
; TELEFAX: (650)328-3377  
; TELEX: N/A  
; INFORMATION FOR SEQ ID NO: 21:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 5620 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: other nucleic acid  
; DESCRIPTION: /desc = "plasmid"  
US-08-892-873-21  
  
Query Match 31.4%; Score 554.4; DB 3; Length 5620;  
Best Local Similarity 59.2%; Pred. No. 2.7e-107;  
Matches 967; Conservative 0; Mismatches 661; Indels 6; Gaps 1;  
  
Qy 2 cagtgcagaactttaggatcaaaaatggagaagaacaacattagcagtgagagcgctcct 61  
Db 1265 CATTCCGGTACTGTGTGTAATAATGGAGACGCCAAACATATAAAGAAAGGCCCGCGGCA 1324  
Qy 62 cgtgatagtccttcctgctgcgcgaggaacaacattataaccaatcattgtataaaattt 121  
Db 1325 TTCTATCTCTAGAGGATGGAAACCGCTGGAGCAACTGCATAGGCTATGAAGAAATAC 1384  
Qy 122 gcatctttctcgaagcaaa-----taatogatgctcotaacaatgaagtaataatcat 175  
Db 1385 GCCCTGGTTCTCGAACAATTTGCTTTTACAGATGCACATATCGAGGTGAACATCACGTAC 1444  
Qy 176 gctcaaatatttgaacacagctgcgcttagctgttagtatagacaacattatgcttgaat 235  
Db 1445 CGGAATACTTTCGAATATGTCGGTTCGGTGGCAGAACGATATGAACCGATATGGGCTGAAT 1504  
Qy 236 gaaacaacatgttgggtgtatgtagtgaacacaataataaaacttttttaactcctgtcctt 295  
Db 1505 ACAAATCACAGAATCGTGTATGCAGTGAAACATCTCTTCAATTTCTTATCCCGGTGTTG 1564  
Qy 296 gctgctttatacttaggaatacccgtaggaacatacaaatgatgatgacacagatggagag 355  
Db 1565 GCGCGTTATTTATCGGAGTTGCAGTTTCGCCCGCGCAACGACATTTTATTAATGAACGTGA 1624  
Qy 356 ttaactggtcatttgaatatatacaaaacaaactatcattgttttagtccaaagaagcactc 415  
Db 1625 TTGCTCAACAGTATGAACATTTCCGACGCTACCGTGTGTTTGTTCCTCAAAAGGGGTTG 1684  
Qy 416 ccgcttattctgagatagacagaaaaactaaagtcttcaaaagtcgtagtattgat 475  
Db 1685 CAAAAATTTTGAAGCTGCAAAAAAATTAACCAATAATCCAGAAAAATTTATATCATGGAT 1744  
Qy 476 agcatgtacgacattaatggcgttgaatgogtatctacacttctgtgcacgttactactgac 535  
Db 1745 TCTAAAACGGATTACAGGGATTTTCAGTCGATGTACAGTTTCGTCACATCTCATCTACCT 1804  
Qy 536 cacacctttagctcattgtcattacacaaaagatattgtatcccttgaaaaaatcgca 595  
Db 1805 CCCGGTTTTAATGATAGGATTTTGTACAGAGTCCCTTTGATCGTGACAAAAAATTTGCA 1864  
Qy 596 ttaattgtcatcatcttggaacaactggattgctgaaggtgtagtagtagagccataga 655  
Db 1865 CTGATAATGAATTCCTCTGGATCTACTGGGTTACTAAGGGTGTGGCCCTTCCGCATAGA 1924  
Qy 656 agtctaactataagattcgttcatacagaaggatcccatattatggcactcgtacaggttcca 715  
Db 1925 ACTGCCCTGCGTCAGATTTCTCGCATGCCAGAGATCCTATTTTTTGGCAATCAAAATCATTCG 1984

Qy 716 caaacatcaattcttctcttagtagaccgttccatcatgctcttggaatttactacatta 775  
Db 1985 GATACGCGATTAAAGTGTGTTCCATTCATCACGCTTTGGAAATGTTTACTACACTC 2044  
Qy 776 tcttactttagtaggacttaaggttgtaattgttgaagaatttgaggcgacacttttc 835  
Db 2045 GGATATTTGATATGGGATTCGAGTCGCTTAATGTATAGATTGGAAGAAGCGTGT 2104  
Qy 836 ttaaaaaacatacagaatacaaaaaatccccactatgttagtggccccctccagttatg 895  
Db 2105 TTACGATCCCTTCAGGATTAACAAATTCAAAAGTGGCTGCTAGTACCAACCCATTTTCA 2164  
Qy 896 ttttggctaaagcccatagtcgatacaatacatttcatcagcttaacgaagattct 955  
Db 2165 TTCCTGCCAAAGACCTCTGATGTGCAAAATACGATTTATCTAATTTACGAAATGCT 2224  
Qy 956 actgagagagctctcttaggaanaagatctcgacagacagtagcaaaaggttgaaatta 1015  
Db 2225 TCTGGGGCGCACCTCTTTTGAAGAAGTCGGGGAAGCGTTGCAAAACGCTTCCATCTT 2284  
Qy 1016 cctggaatcatacagaagatatgatttaactgaaccttgcgtcgtgtaattgattaccct 1075  
Db 2285 CCAGGGATACGACAAAGGATATGGGCTCCTCAGAGACTACATCAGCTATTCTGATTACACC 2344  
Qy 1076 cataatgctlgaaacacaggttcaactggaagacccttgcatacattaaagctaaagt 1135  
Db 2345 GAGGGGATGATAACCGGGCGGCTCGGTAAAGTTGTTCCATTTTGAACGGAAGGTT 2404  
Qy 1136 ttagataacgctctgggaagcgctagaccagagagaagcggaataatgctttcaa 1195  
Db 2405 GTGGATCTGGATACCGGAAACGCTGGCGTTAATCAGAGCGGCAATATGTGTCAGA 2464  
Qy 1196 agtgaatgattatgaagggatattacacaacatccggaagcaactattgatactattgac 1255  
Db 2465 CGACCTATGATTATCTCGGTTATGTAAACAATCCGAAGCGACCAACGCCCTTGATTGAC 2524  
Qy 1256 aaagatgttggtctcattctggagatatgtgattacgacgaagatggaattcttt 1315  
Db 2525 AAGGATGGATGGCTACATTCGGAGACATAGCTTACTGGGACGAAGACACATCTCTTC 2584  
Qy 1316 atagttgctgattgaagaacattattaaatacagaaggatatacaggttgcgctctgaa 1375  
Db 2585 ATAGTTGACCGCTTCAAGCTTTTAATTAATAACAAGGATATCAGTGGCCCCCTCGAA 2644  
Qy 1376 ctggaanaatctgcttttacaacatccaagtatctgctgtagtgggttactggaagttccg 1435  
Db 2645 TTGGAATCGATATTGTTACACACCCCAACATCTTCGACGCGGGCGTGCAGGTCTTCCC 2704  
Qy 1436 gacgaatttgggacaattaccctgctgtgtgtgtgttagaattcggcaagacgtg 1495  
Db 2705 GACGATGACCGCGGTGAACCTTCCCGCCCGCTGTGTGTTTGGAGCACGGAAAGACGATG 2764  
Qy 1496 actgaaaaggaagtccaagatttattcagcacagtcactccaacaaagcatcttca 1555  
Db 2765 ACGGAAAAGAGATCGTGGATTACCTCGCCAGCTCAAGTAACNACCCGCAAAAAGTTGCGC 2824  
Qy 1556 ggcggtgctgattttagacagtagtattccgaagggccctacttgaaactcatcagaag 1615  
Db 2825 GGAGGAGTGTGTTGTGACGACGAAGTACCGAAAGGTCTTACCGGAAAACTCGACGCAAGA 2884  
Qy 1616 gagctcccgagaaat 1629  
Db 2885 AAAATCAGAGAGAT 2898

## RESULT 6

US-09-334-765A-21

; Sequence 21, Application US/09334765A  
; Patent No. 6238893  
; GENERAL INFORMATION:  
; APPLICANT: Fallaux, Fris J.  
; APPLICANT: Hoebein, Robert C.

; APPLICANT: Bout, Abraham  
; APPLICANT: Valerio, Domenico  
; APPLICANT: Van der Eb, Alex J.  
; TITLE OF INVENTION: PACKAGING SYSTEMS FOR HUMAN RECOMBINANT ADENOVIRUS TO BE USED  
; TITLE OF INVENTION: IN GENE THERAPY  
; FILE REFERENCE: 3833.2US  
; CURRENT APPLICATION NUMBER: US/09/334,765A  
; CURRENT FILING DATE: 1999-06-16  
; PRIOR APPLICATION NUMBER: US 08/793,170  
; PRIOR FILING DATE: 1997-03-25  
; PRIOR APPLICATION NUMBER: PCT/NL96/00244  
; PRIOR FILING DATE: 1996-06-14  
; PRIOR APPLICATION NUMBER: EP 95201728.3  
; PRIOR FILING DATE: 1995-06-26  
; PRIOR APPLICATION NUMBER: EP 95201611.1  
; PRIOR FILING DATE: 1995-06-15  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: Corel WordPerfect 8.0  
; SEQ ID NO 21  
; LENGTH: 5620  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; NAME/KEY: Ad5 left terminus  
; LOCATION: 1..457  
; NAME/KEY: enhancer  
; LOCATION: 458..969  
; NAME/KEY: exon  
; LOCATION: 970..1204  
; NAME/KEY: gene  
; LOCATION: 1218..2987  
; NAME/KEY: polyA\_signal  
; LOCATION: 3018..3131  
; NAME/KEY: pUC12 backbone  
; LOCATION: 3132..5620  
; NAME/KEY: gene  
; LOCATION: 4337..5191  
; OTHER INFORMATION: Description of Artificial Sequence: Plasmid pICL  
US-09-334-765A-21

Query Match 31.4%; Score 554.4; DB 4; Length 5620;  
Best Local Similarity 59.2%; Pred. No. 2.7e-107;  
Matches 967; Conservative 0; Mismatches 661; Indels 6; Gaps 1;

Qy 2 cagtgcaagacttttagggatcaaaaatggaagaagaaaacatttaggcagtgagagcgtcct 61  
Db 1265 cattccggtactgttggtaaaatggagacgccaataaagaaagcccgccgcca 1324  
Qy 62 cgtgatatagtcctcctggctggcgagacacaaattatataccaatcattgtataaaatt 121  
Db 1325 ttctatcctctagagatggaaacgctggagcgaactgcataagggctatgaagaaatac 1384  
Qy 122 gcattcttctcgaagcaa-----taatcgatgctctacacaaatgaagtaatacatc 175  
Db 1385 gccctggttctcgaacaaattgcttttcagatgcacataatcgaggtgaacatacgtac 1444  
Qy 176 gctcaaatatttgaacacagctgcgccttagcttagtatagaacaatatggcttgaat 235  
Db 1445 gcggaatacttcgaaatgctcgttgcgttggcagaagctatgaacgatatggcgtgaat 1504  
Qy 236 gaaaaacaatgttgggtgtatgcagtgaaaaacaataaaacttttttaactcgtcctt 295  
Db 1505 acaaatcacagaatcgtgtagtcgagtgaataactctcttcaattcttctgctggtgtg 1564  
Qy 296 gctgcttttacttaggaatacaccagtagcaacatcaaatgatatatcacagatgagag 355  
Db 1565 ggcggttattatcgagggtgcagttgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1624  
Qy 356 ttaactggttcatttgaataatatacaaaacaaactatcatgttttagttcaagaagaaac 415  
Db 1625 ttgctcaacagtatgaacatttcgcagcctaccgtagtggtttgttttcccaaaaggggtg 1684



Qy	356	ttactggttcatttgaaataatacaaaaacaaactatcatgttttagttcaagagaagcaactc	4135
Dy	1625	ttctcaacagatgaaacttccgcagcctacccgtagttgtttgtttccaaaaggggttg	1684
Qy	416	ccgtctattctgagatcacagcaaaatcaaagttctataaaaaagtcgtagttatcgat	475
Dy	1685	caaaaaatttgaaacgtgcnaaaaaaattaccaaataaccagaaaaattattatcatcgat	1744
Qy	476	agcatgtacagacattaatggcgttgaaatcgctatctacctttgttgacgttatactgac	535
Dy	1745	tctaaaacggattaccagggatttcagtcogtgcacagttcgicacatctcatctacct	1804
Qy	536	caacctcttgatccattgctcatttacacaaaagattttgaccccttgaaaaaatcgca	595
Dy	1805	cccggttttaatgaaacagattttgaccagagtcctttgatctgacaaaacaattcgca	1864
Qy	596	ttaatattgtcatctatcgaaacaactggattgcttaagggtgtagttactgagccataga	655
Dy	1865	ctgataatgaattccctctggtactactggtttacctaagggtggtgccccttcgcataga	1924
Qy	656	agttctaactaagatttgcttatagcagggatcccatatttatggcaactcgtacaggttcca	715
Dy	1925	actgctgctcgatattctcgatccgagatgcctatttttggcaatcaaatcatccg	1984
Qy	716	caaaatcaaatcttcccttagctacogttcccatcgcttttggaaatgtttactacatta	775
Dy	1985	gatactgcgattttagtgtgtctcatcccatcaggttttgaaatgtttactacactc	2044
Qy	776	tcttacttttagtaggaacttaaggttgttaatgttgaagaaaattttgaggcgacttttc	835
Dy	2045	ggatatttgatatgtggaatttgagtcgtcttaatgtatagatttgaagaagagctgttt	2104
Qy	836	ttaaaaaccatacagaattataaaaaatcccaactatgttagtggccctccagttatggtg	895
Dy	2105	ttacgatccctcaggtattacaaaattcaaatgctgtgttagtaccacccctatttcca	2164
Qy	896	tttttgctaaagcccattagtcgatcaatacagattttatcgagcttaacggaagttgct	955
Dy	2165	tcttcgcaaaagcaactgtatgcaaatcacgatttatctaatctcacggaatttgct	2224
Qy	956	actggagagctccttttagaaaaagatgctgcagagcagtagcaaaagagtttgaataa	1015
Dy	2225	tctggggcgccactcttctcgaaaagtcgggggaagcggttgcaaaacgcgtccactct	2284
Qy	1016	cctggaatcatacaaggatatggattaaactgaaacttgcgctgttaatgattaccacct	1075
Dy	2285	ccagggtacgacaaggatatgggtccactgagatcatcacgtctattctgattacaccc	2344
Qy	1076	cataatgctgtgaacaggttcaactggaaagcccttgccatacatataaagctaaaagt	1135
Dy	2345	gagggggatgataaaacggcgcgctggtaaatgtgttccatttttggaaagcgaaggtt	2404
Qy	1136	ttagataacgtactctgggaagcgcttaggcaccagagaaaagcggcaaatatgtcttcaa	1195
Dy	2405	gtggtatcgataccgggaacacgcctggcggttaatcacagagggcgaaattatgtgtcaga	2464
Qy	1196	agtgaatgattatgaaggaattatacaacaatccggaagcaactattgatactattgac	1255
Dy	2465	ggacctatgatattgcgggttatgttaacaacatccggaagcgcaacgcgtttagtgc	2524
Qy	1256	aaagatggttgcttcatctcgagatatggatatattcacgcagaagatggaaaattctttt	1315
Dy	2525	aaggatggaatggctattctggagacatagcttactgggacgaagacacactcttctc	2584
Qy	1316	atagttgatcgattgaagaacttataaatacagaaggatatacaggttgcgcctgctgaa	1375
Dy	2585	atagttgacgcgtgaagtccttaattaaatacaaggaatatcaggtggcccgccgtgaa	2644
Qy	1376	ctggaaaaatctcttttacaaacatccaagtattgtgtatgvgggtgttacttggagttccg	1435
Dy	2645	ttggaaatcgatattgttatacaaccccacactcttcgacggcggtgtggaggttttccc	2704

```

Qy 1436 gacgaatttggtggacaataaccctgcttgcctttgttgtttagaaatcctggcaaacgctg 1495
Db 2705 gagcatgacctcggttaacttcctccgcgcgtttgttttggagcacggaaaacgatg 2764
Qy 1496 actgaaaaagaagtccaagatttttatgcagcacaaagtcactcccacaacaaagcatcttcga 1555
Db 2765 acgaaaaaagagatcgtggattacgtcgccagtcaagttaaacacocgcgaaaaaggttgcg 2824
Qy 1556 ggoggtgtcgtatttttagacagtagtattccgaaagccctactcctaagaaactcatacgaagaag 1615
Db 2825 ggagaggttggtttgtggacgaagaccgaaaggtcttacgcgaaaactcgacgcgaaga 2884
Qy 1616 gagctccagagaat 1629
Db 2885 aaatcacagagat 2898

RESULT      8
US-09-333-820-21
; Sequence 21, Application US/09333820A
; Patent No. 6306652
; GENERAL INFORMATION:
; APPLICANT: Fallaux, Frits J.
; APPLICANT: Hoebein, Robert C.
; APPLICANT: Bout, Abraham
; APPLICANT: Valerio, Domenico
; APPLICANT: Van der Eb, Alex J.
; TITLE OF INVENTION: PACKAGING SYSTEMS FOR HUMAN RECOMBINANT ADENOVIRUS T
; TITLE OF INVENTION: GENE THERAPY
; FILE REFERENCE: 3833 IUS
; CURRENT APPLICATION NUMBER: US/09/333,820A
; CURRENT FILING DATE: 1999-06-15
; EARLIER APPLICATION NUMBER: US 08/793,170
; EARLIER FILING DATE: 1997-03-25
; EARLIER APPLICATION NUMBER: PCT/NL96/00244
; EARLIER FILING DATE: 1996-06-14
; EARLIER APPLICATION NUMBER: EP 95201728.3
; EARLIER FILING DATE: 1995-06-26
; EARLIER APPLICATION NUMBER: EP 95201611.1
; EARLIER FILING DATE: 1995-06-15
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Corel WordPerfect 8.0
; SEQ ID NO 21
; LENGTH: 5620
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: Ad5 left terminus
; LOCATION: 1..457
; NAME/KEY: enhancer
; LOCATION: 458..969
; NAME/KEY: exon
; LOCATION: 970..1204
; NAME/KEY: gene
; LOCATION: 1218...2987
; NAME/KEY: polyA_signal
; LOCATION: 3018..3131
; NAME/KEY: pUC12 backbone
; LOCATION: 3132..5620
; NAME/KEY: gene
; LOCATION: 4337...5191
; OTHER INFORMATION: Description of Artificial Sequence: Plasmid pICL
US-09-333-820-21
```

```

Query Match      31.4%  Score 554.4;  DB 4;  Length 5620;
Best Local Similarity 59.2%;  Mismat.No. 2.7e-107;
Matches 967;  Conservative 0;  Pmidmatch 661;  Indels 6;  Gaps 1;
Qy      2  cagtgcaagactttagggtatcaaaatggaaagaagaaacattaggcatggagagcgctct 61
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      1265  cattccgctactgttggtaaatgaaacacgcaaaatacaataaaagaaagccgcgcgcga 1324

```

[illegible]

## RESULTS

RESULTS  
US-08-867-352-22

US-08-067-332-22  
: Sequence 22. Application US/088673352

: Patent No. 6060273

; PALCIL NO. 6060273  
: GENERAL INFORMATION:

GENERAL INFORMATION:  
APPLICANT:

;  
; AFFILIANT:  
; TITLE OF INVENTION: Multicistronic expression units and their use

; TITLE OF INVENTION: MULTIPLE SEQUENCE ALIGNMENT  
 ; NUMBER OF SEQUENCES: 25

; NUMBER OF SEQUENCES: 2  
: COMPUTER READABLE FORM:

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```

COMPUTER: IBM PC compatible

COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPA)

```

; SOFTWARE:  FALCON III RELEASE 1.0
;
; CURRENT APPLICATION DATA:

```

; CURRENT APPLICATION DATA: US/08/867.352  
 ; APPLICATION NUMBER:

APPLICANT :  
FILING DATE:

CLASSIFICATION:

CLASSIFICATION:  
PRIOR APPLICATION DATA:

; PRIOR APPLICATION DATA: 08/387.847  
 ; APPLICATION NUMBER: 08/387.847

APPLICANT  
FILING DATE

; FILING DATE: 22:  
; INFORMATION FOR SEQ ID NO: 22:; INFORMATION FOR SEQ ID NO:  
: SEQUENCE CHARACTERISTICS:

SEQUENCE CHARACTERISTICS:  
LENGTH: 1811 base pairs

```

; LENGTH: 1811 base pairs
; TYPE: nucleic acid

```

STRANDEDNESS: single

TOPOLOGY: Linear

MOLECULE TYPE: CDNA

MOLECULE TYPE:  
ORIGINAL SOURCE:

ORGANISM: fire fly (photinus pyralis)

IMMEDIATE SOURCE:

CLONE: PRSVLUC (de Wet et al., 1987)

FEATURE:

NAME/KEY: CDS

LOCATION: 94..1743

```

; LOCATION: 34.174
; OTHER INFORMATION: /note= "coding region of the

```

OTHER INFORMATION: luciferase gene; flanked by 5' -SmaI and

OTHER INFORMATION: 3'-HindIII restriction cleavage sites"

RESULT 10  
US-08-862-431-32  
: Sequence 32, Application US/08862431  
: Patent No. 6120994  
: GENERAL INFORMATION:  
: APPLICANT: TAM, SHUI-PANG  
: TITLE OF INVENTION: ANTIOXIDANT RESPONSIVE ELEMENT  
: NUMBER OF SEQUENCES: 51  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
: STREET: 1100 NEW YORK AVENUE, SUITE 600  
: CITY: WASHINGTON  
: STATE: DC  
: COUNTRY: US  
: ZIP: 20005-3934  
: COMPUTER READABLE FORM:

;; COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/862,431  
FILING DATE: 23-MAY-1997  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Kim, Judith U.  
REGISTRATION NUMBER: 40,679  
REFERENCE/DOCKET NUMBER: 1669.0020000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 371-2600  
TELEFAX: (202) 371-2540  
INFORMATION FOR SEQ ID NO: 32:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5789 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: circular  
US-08-862-431-32

Query Match 31.3%; Score 552.8; DB 3; Length 5789;  
Best Local Similarity 59.1%; Pred. No. 5.8e-107;  
Matches 966; Conservative 0; Mismatches 662; Indels 6; Gaps 1;

```
Qy 2 cagtcgaagacttaggatcaaaatggaagaagaacatttaggcattgagagcgctct 61
Db 247 CATTCGGGTACTGTGGTAAATGGAAGACGCCAAACATAAAGAGGCCCGCGCCA 306
Qy 62 cgtgatatagtcacactcctggctcgcgaggaacaacattatccaatcattgtataaatt 121
Db 307 TCTATCTCTAGAGGATGGAACCGCTCGAGAGCAACTGCTAAGGCTATGAAGAGATAC 366
Qy 122 gcattctttctgaagcaa-----taacgatgctacatacaaaatgaagtaatatcat 175
Db 367 GCGCTGGTTCCTGGAAACAATTGCTTTACAGATGCACATATCGAGGTGAACATCACGTAC 426
Qy 176 gctcaaatattgaaacagctgcgcttagctgttagtatagaaacaatatagcttgat 235
Db 427 GCGGAATACTTCGAAATGTCGGTTGCGAGAAAGTATGAAACGATATGGGCTGAAT 486
Qy 236 gaaacaactgttggtggtgtagtcagtgaaacaacataaaactttttaaactctgctct 295
Db 487 ACAATACAGAACTGCTGATGCACTGAAACACTCTCTCAATCTTTATGCCGTGTG 546
Qy 296 gctgcttatacttaggaataaccagtagcaacaacatacaatgatgatgacacagatggag 355
Db 547 GCGCGTTATTATCGGAGTTGCACTTGGCGCGGCAACGACATTTAATGAACGTGAA 606
Qy 356 ttaactggtcattgaaatatcaaaacaaactatcagtttagttcaaaagaagcactc 415
Db 607 TTGCTCAACAGATGAACATTTGCGAGCTACCGTAGTGTGTTTCCAAAAGGGGTG 666
Qy 416 ccgcttattcagagtagcacgaaactaagtttcaataaaagtcgtagttatcgat 475
Db 667 CAAAAAATTTGAAGCTGCAAAAAAATTAACCAATAATCCAGAAAAATTTATATCATGAT 726
Qy 476 agcatgacacattaaatggcgtgaaatgcgtatatacttctgtgacgttatactgac 535
Db 727 TCTAAACGGATTACAGGGATTTTCAGTCGATGTCACGTTGCTCAGATCTCATCTACCT 786
Qy 536 cacaccttgatccattgcatcaccacaaagaattttgatcccttgaaaaaatcgca 595
Db 787 CCCGGTTTAATGAATACGATTTGTACCCAGATCTTTGATCGTGCACAAACAAATTGCA 846
Qy 596 ttaattatgcatcatcattggaacactgattgcctaagggtgtagtactgagccataga 655
Db 847 CTGATAATGAATTCCTCTGGATCTACTGGGTACCTAAGGGTGTGGCCCTTCCCGATAGA 906
Qy 656 agtctaactaagaattcgttcatagcaggggaggtccatttatggcaactcgtacggtcca 715
```

```
Db 907 ACTGCTCGTCAGATTCTCGCATGCCAGATCCTATTTTGGCAATCAATCATTCGCG 966
Qy 716 caaacataaattcttcttagtagccggtccatcatgcttgcgaattgttactacatta 775
Db 967 GATACCTCGATTTTAAAGTGTGTTCCATTCCTACCGGTTTGGAAATTTTACTACATC 1026
Qy 776 tcttactttgtagtaggaacttaaggttgaatgttggaagaattttgagggcgcaatttc 835
Db 1027 GGATATTGATATGTGGATTTCGAGTCCTTAATGTATAGATTGAAGAAGAGCTGTT 1086
Qy 836 ttaaaacacatacagaattacaaaatccccactattgttagtggccccctccagttgat 895
Db 1087 TTACGATCCCTTCAGGATTACAAAATTCAAAAGTGGTGTAGTACCAACCTATTTCA 1146
Qy 896 ttttggctaaaagcccattagtgcatacagatttatcgagcttaacagaagttgct 955
Db 1147 TTTCTGCGCAAAAGCACCTCTGATTGACAAATACGATTTATCTAATTTACAGAAATGCT 1206
Qy 956 actggagagctccttttaggaaaagatgctcagagaagcagtagcaaaagggtgaaatta 1015
Db 1207 TCTGGGGCGCACCTCTTTTCGAAAGAACTCGGGGAAGCGTTGCAAAACGCTTCCATCT 1266
Qy 1016 cctggaatcatacaaggatatggattaaactgaaacttgcgtgctgtaattgattaccct 1075
Db 1267 CCAGGGATACGACAAAGGATATGGGCTCCTAGACTACATCAGCTATTCTGTATTACACC 1326
Qy 1076 cataatgctgigaacaaggttcaactggaagacccttgccatcacattaaagcttaagtt 1135
Db 1327 GAGGGGATGATAAACCGGCGCGTCCGTAAAGTGTGTTCCATTTTGAAGCGAAGGT 1386
Qy 1136 ttagataacgctactctgggaagcgctaggaccagaggaagggcgaaatatgctttcaa 1195
Db 1387 GTGGATCTGGATACCGGGAACACGCTGGCGTTAATACAGAGAGGGAATTTATGTGTCAGA 1446
Qy 1196 agtgaatgatatgaaaggatatatacaacaatccoggaagcaactatgatactattgac 1255
Db 1447 GGACCTATGATTATGTCGGTATATGTAACAATTCGGAAGGACCAACGCGCTTGATTGAC 1506
Qy 1256 aaagatggttggtctcattcggagatatggaattatcacgaagatggaaattctctt 1315
Db 1507 AAGGATGATGGCTACATTTCTGGAGACATAGCTTAACTGGGACGAAGACGAACTTCTTC 1566
Qy 1316 atagtgcctgattgaagaacttatataacaaaggatatcaggttgcgctgctgaa 1375
Db 1567 ATAGTTGACCGCTTGAAGCTCTTAATTAATACAAAGGATATCAGTGGCCCCGCTGAA 1626
Qy 1376 ctgaaaaatctgcttttacaacatccaagtattctgctgtagtgcgggtgctactggagtc 1435
Db 1627 TTGGAATCGATATTGTTACAAACACCCCAACATCTTTCGACGGGGGTGGCAGGTCTCC 1686
Qy 1436 gacgaatttgtagacaattacctgctgctgctgctgctgctgctgctgctgctgctg 1495
Db 1687 GACGATACGCGGTGAACTTCCCGCCCGCTGTTGTTGTTGGACGCGGAAGACGATG 1746
Qy 1496 actgaaaggagttcaagattttattgcagcacaagtcactcccaacaaagcattctga 1555
Db 1747 ACGGAAAAGAGATCGTGGATTACCTGCCAGTCAAGTTAACACCGCGCAAAAAGTTGGC 1806
Qy 1556 ggcgggtgctgattgttagcagattatccgaaagggccctactggaactactcagaag 1615
Db 1807 GGAGGAGTGTGTTTGTGGAGGAGTACCGAAAGGCTTTACCGGAAAACTCGACGCAAGA 1866
Qy 1616 gagctccgagaaat 1629
Db 1867 AAAATCAGAGAGAT 1880
```

RESULT 11  
US-08-862-431-31  
; Sequence 31, Application US/08862431  
; Patent No. 6120994  
; GENERAL INFORMATION:



APPLICANT: TAM, SHUI-PANG  
 TITLE OF INVENTION: ANTIOXIDANT RESPONSIVE ELEMENT  
 NUMBER OF SEQUENCES: 51  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
 STREET: 1100 NEW YORK AVENUE, SUITE 600  
 CITY: WASHINGTON  
 STATE: DC  
 COUNTRY: US  
 ZIP: 20005-3934  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release v1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/862,431  
 FILING DATE: 23-MAY-1997  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Kim, Judith U.  
 REGISTRATION NUMBER: 40,679  
 REFERENCE/DOCKET NUMBER: 1669.0020000  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (202) 371-2600  
 TELEFAX: (202) 371-2540  
 INFORMATION FOR SEQ ID NO: 31:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 5791 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: circular  
 US-08-862-431-31

	Query Match	31.3%	Score 552.8;	DB 3;	Length 5791;
	Best Local Similarity	59.1%;	Pred. No. 5.8e-107;		
	Matches	966;	Conservative	0;	Mismatches 662; Indels 6; Gaps 1.
<hr/>					
Qy	2	cagtgcaagaactttagggatcaaaatgggaagaagaanaaacatttaggcattgagagcgctcct	61		
Db	249	CATTCCGGTACTGTGTPAAAATGGAGACGCCAATAAAACATATAAGAAAGAGCCCGCGCCA	308		
Qy	62	cgtagatatagccacccctcggtcggcagagcaacaattataccaatcatcttgtataaaattt	121		
Db	309	TTCTATCCTTAGAGGATGGAAACCCTGGAGCAACTGCATTAAGGCTATGAAGAGATAC	368		
Qy	122	gcattcttttcctgaagcaa-----taactgatgctcatacaaaaatgaagtaatacatat	175		
Db	369	GCCCTGGTTCCTGGAACAATTGCTTTTTACAGATGCACATATCGAGGTGNACATCACGTAC	428		
Qy	176	gctcaaatatttgaaaccagctgcgcttagctgttagtagaacaatatgcttgaat	235		
Db	429	CGGGAATACTTTCGAATAATCGCTTGCGTTGGCAGAAGCATGAAACGATATGGGCTGAAT	488		
Qy	236	gaaaaacaatgttggggtgatgcagtgaataaacaataaaaactttttaactcctgcctt	295		
Db	489	ACAAATCACAAATCGTCGTATGCAGTCAAACCTCTTCAATTTCTTATGCGCGTGTG	548		
Qy	296	gctgctttactactlaggaataccagtagcaacatacaatgatatatcacagatgaagaag	355		
Db	549	GGCGCGTTATTTATCGGAGTTCAGTTCGCCCGCGGAACGACATTTATTAAGAACGTGAA	608		
Qy	356	ttaaactggctatttgaatatatacaaaaccaactatcatgttttagttcaagaagaacactc	415		
Db	609	TTGCTCAACAGTATGAACATTTCCGCAGCCTACCGTAGTGTGTTTCCAAAAAGGGGTG	668		
Qy	416	ccgcttatcttgagtagtacagcaaaatctaagtttcattataaaaaagtcgtagttatcgat	475		
Db	669	CAAAAAATTTTGAAGTCGCAAAAAAATTACCAATAATCCAGAAATATTATTATCATGGAT	728		
Qy	476	agcatgtacgacattaaatggcgcttgaatgcgatcatcactctttgttgcacgttatactgac	535		

Qy 1616 gagctccgagaaat 1629  
| | | | |  
Db 1869 AAAATCAGAGAT 1882

RESULT 12  
US-08-862-431-29  
: Sequence 29, Application US/08862431  
: Patent No. 6120994  
: GENERAL INFORMATION:  
: APPLICANT: TAM, SHUI-PANG  
: TITLE OF INVENTION: ANTIOXIDANT RESPONSIVE ELEMENT  
: NUMBER OF SEQUENCES: 51  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: STERNER, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
: STREET: 1100 NEW YORK AVENUE, SUITE 600  
: CITY: WASHINGTON  
: STATE: DC  
: COUNTRY: US  
: ZIP: 20005-3934  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: PatentIn Release #1.0, Version #1.30  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/08/862,431  
: FILING DATE: 23-MAY-1997  
: CLASSIFICATION: 435  
: ATTORNEY/AGENT INFORMATION:  
: NAME: Kim, Judith U.  
: REGISTRATION NUMBER: 40,679  
: REFERENCE/DOCKET NUMBER: 1669.0020000  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: (202) 371-2600  
: TELEFAX: (202) 371-2540  
: INFORMATION FOR SEQ ID NO: 29:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 5793 base pairs  
: TYPE: nucleic acid  
: STRANDEDNESS: double  
: TOPOLOGY: circular  
US-08-862-431-29

Query Match 31.3%; Score 552.8; DB 3; Length 5793;  
Best Local Similarity 59.1%; Pred. No. 5.8e-107;  
Matches 966; Conservative 0; Mismatches 662; Indels 6; Gaps 1;

Qy 2 cagtgcaagactttaggatcaaaatggaagaagaaacattagcatggagagcgtcct 61  
| | | | |  
Db 251 CATTCGGTACTGTGGTAAAATGGAACACGCCCAAAACATAAAGAGCCCGGCCCA 310

Qy 62 cgtgatagtcacacctcgtggtcgccaggaacaacaattataccatcatctgtataaatt 121  
| | | | |  
Db 311 TTCTATCTCTAGAGGATGGAACCGCTGGAGCAACTGCATAGGCTATGAAGAGATAC 370

Qy 122 gcattcttctgaagca-----taactgctctacatacaaatgaagtaatacatat 175  
| | | | |  
Db 371 GCCCTGGTTCTTGGAAACAAATTCGCTTTACAGATGCACATATCGAGGTGAAACATCACGTAC 430

Qy 176 gctcaaatatttgaacaccagctgcgccttagctgttagtatagaacaataatggcttgaat 235  
| | | | |  
Db 431 GCGGAATACTTCGAAATGCTCCGTTGGCAGAAGCTATGAAACGATATGGGCTGAAT 490

Qy 236 gaaaacaattgtgtggtgtatgatgaagtgaacaacaataaaactttttaaactctgtcctt 295  
| | | | |  
Db 491 ACAAAATCACAAATCGCTGTATGCAAGTGAAGAAACTCTCTTCAATTTATGCCGGTGTG 550

Qy 296 gctgcttttatacttaggataaccagtagcaacatacaaatgatatgtacacagatggagag 355  
| | | | |  
Db 551 GCGCGGTATTATTTCGGAGTTGTCAGTTGCGCGCGGCAACGACATTTAATGAACGTGAA 610

Qy 356 ttaactggtcatttgaatatatacaaaacaaactatcatgttttagttcaaaagaagcactc 415  
| | | | |  
Db 611 TTGCTCAACAGTATGAACATTTTCGACGCTACCGTAGTGTGTGTTTCCAAAAGGGGTG 670

Qy 416 ccgcttattctgagagtacagcaaaatactaaagtttcaataaaaaagtcgtagtattcgat 475  
| | | | |  
Db 671 CAAAAATTTTGAACGTGCAAAAAAATTTACCAATAATTTCCAGAAAATTTATTTATCATGGAT 730

Qy 476 agcatgtacgacattaatgcggtgaatgcgtatctactactcttctgcaactgttaactgac 535  
| | | | |  
Db 731 TCTAAAACGGATTTACAGGGATTTTCAGTCGATGTACAGGTTCTGTCACATCTCATCTACCT 790

Qy 536 cacacctttgatccattgtctattcacacaaaagattttgatccctcttgaaaaaactcgca 595  
| | | | |  
Db 791 CCCGGTTTAAATGAATACGATTTTGTACACAGATCCCTTTGTATCGTGCACAAAACAATTCGA 850

Qy 596 ttaattatgcatcatctgtgaaacactggattgctcaaggtgtagtactgagcagcataga 655  
| | | | |  
Db 851 CTGATAATGAATTCCTCTGGATCTACTGGTTACTTAAGGGTGTGGCCTTTCCGCATAGA 910

Qy 656 agtctaactataagattcgttcatagcagggatccccatttatggcactcgtacggttcca 715  
| | | | |  
Db 911 ACTGCTGCGTCAGATTCGCAATGCCAGATCCTATTTTGGCAATCAAAATCATTCCTCG 970

Qy 716 caaacatcaattcttcttagtaccgttccatcatgctcttggaaatgtttactacattta 775  
| | | | |  
Db 971 GATACGCGAATTTAAAGTGTGTTCCATTCATCAGGGTTTGGAAATGTTTACTTACACTC 1030

Qy 776 tcttacttggtagtaggacttaaggttgtaattgtggaagaaatttgagggcgcaacttttc 835  
| | | | |  
Db 1031 GGATATTTGATATGTGGATTTTCGAGTCTCTTAATGTATAGATTTGAAGAAGAGAGTGTT 1090

Qy 836 ttaaaaccatacagaattacaaatcccactattgtagtgccctcccagtttatggtg 895  
| | | | |  
Db 1091 TTACGATCCCTTCAGGATTACAAATTCAAAGTGGTTCGTAGTACCAACCCCTATTTTCA 1150

Qy 896 ttttggctaaaagccatttagtcgatcaatcacatgtattatcgagcttaaacggaagtgc 955  
| | | | |  
Db 1151 TTCCTCGCAAAAGCACCTCTGATTGACAAATACGATTTATCTAATTTACACGAAATTCGT 1210

Qy 956 actggaggagctcctttaggaaagagatgctgcagaagcagtagcaagagagttgaaatta 1015  
| | | | |  
Db 1211 TCTGGGGCGCACCTCTTTTCGAAAAGAGTGGGGAAGCGTTGCAAAACGCTTCATCTT 1270

Qy 1016 cctggaatcatcaaggatagtaactgaactgaactcgtcgtgtaattgattaccct 1075  
| | | | |  
Db 1271 CCAGGGATACCAAGAGATATGGGCTCACTGAGACTAGACTAGCTATTTCTGATTTACACCC 1330

Qy 1076 cataatgctgtgaacaaggttcaactggaagaccccttgccatcacattaaagcttaagtt 1135  
| | | | |  
Db 1331 GAGGGGATGATAAACCGCGCGCTCGGTAAAGTTGTTCATTTTGTGAAGCGAAGGTT 1390

Qy 1136 ttagataacgctactgtggaagcgctaggaacaggaagaaagggcgaataatgctttcaa 1195  
| | | | |  
Db 1391 GTGATCTGSAATACCGGAAAACGCTGGCGTTAATCAGAGAGCGGAAATTTATGTCAGA 1450

Qy 1196 agtgaatgatatgaagagattatcacaaatccggaagcaactattgatctactatgac 1255  
| | | | |  
Db 1451 GGACCTATGATTTATGTCGGTTATGTAACAATTCGGGAAGCGACCAACGCTTGAATTGAC 1510

Qy 1256 aagatggttggcttcttctggagatatgtgatatcagcgaagatggaaaattttttt 1315  
| | | | |  
Db 1511 AAGGATGATGGCTACATTTCTGGACATAGCTTACTGGGACGAAGACGAACACTTCTTC 1570

Qy 1316 atagttgatcagatgaagaacttattaaatacaagggatcaggttgccgctcgtgaa 1375  
| | | | |  
Db 1571 ATAGTTACCGCTTGAAGCTTTAATTAATAACAAAGGATATCAGGTGGCCCCCGCTGAA 1630

Qy 1376 ctggaataatcgtcttttacaacatccaagtatgtgtagtcggtggtggttactggtggtccg 1435  
| | | | |  
Db 1631 TTGGAATCGATATTGTTACACACCCCAACATCTTCGACGCGGGCGTGGCAGGTCTTCCC 1690









GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OW nucleic - nucleic search, using sw model

Run on: September 7, 2002, 19:55:51 ; Search time 296.36 Seconds  
(without alignments)  
10225.246 Million cell updates/sec

Title: us-09-993-874-1  
Perfect score: 1765  
Sequence: 1 tcagtgaagactttaggga.....aaaaaaaaaaaaaaaaaaaaa 1765

Scoring table:  
IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 347872

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Genes032802.\*  
1: /SID55/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SID55/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SID55/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SID55/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SID55/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SID55/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SID55/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SID55/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SID55/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SID55/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SID55/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SID55/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SID55/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SID55/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SID55/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SID55/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SID55/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SID55/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SID55/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SID55/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SID55/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
c 1	575.2	32.6	2193	22	AAF24923 Complement of Fc e
2	575.2	32.6	2268	22	AAF24920 Nucleotide sequenc
c 3	575.2	32.6	2268	22	AAF24921 Complement of Fc e
4	573.8	32.5	1638	22	AAF24919 Nucleotide sequenc
5	571.6	32.4	1642	20	AA32854 Mutant luciferase
6	571.6	32.4	1642	22	AA32854 Mutant luciferase
7	570	32.3	1639	20	AA32847 Mutant luciferase
8	568.4	32.2	1639	20	AA32844 Mutant luciferase
9	566.8	32.1	1639	20	AA32843 Mutant luciferase

10	566.8	32.1	1639	20	AA32845	Mutant luciferase
11	566.8	32.1	1639	22	AA32845	Beetle cDNA encodi
12	565.2	32.0	1639	20	AA32846	Mutant luciferase
13	565.2	32.0	1639	22	AA32846	Beetle cDNA encodi
14	565.2	32.0	1639	22	AA32847	Beetle cDNA encodi
15	563.6	31.9	1639	22	AA32848	Beetle cDNA encodi
16	562.4	31.9	1639	20	AA32849	Mutant luciferase
17	560.8	31.8	1639	20	AA32852	Mutant luciferase
18	560.4	31.8	1639	22	AA32853	Beetle cDNA encodi
19	560.4	31.8	1639	22	AA32854	Beetle cDNA encodi
20	559.8	31.7	1639	16	AA32855	Vaccinia virus vec
21	559.2	31.7	1639	20	AA32856	Mutant luciferase
22	558.2	31.7	1639	22	AA32857	Beetle cDNA encodi
23	558.8	31.7	1639	22	AA32858	Beetle cDNA encodi
24	557.6	31.6	1639	20	AA32859	Mutant luciferase
25	557.6	31.6	1639	20	AA32860	Mutant luciferase
26	556	31.5	1639	22	AA32861	Beetle cDNA encodi
27	555.6	31.5	1639	22	AA32862	Beetle cDNA encodi
28	552.8	31.3	1798	20	AA32863	Firefly luciferase
29	552.8	31.3	1811	15	AA32864	Encodes firefly lu
30	552.8	31.3	5620	18	AA32865	Ad5-ITR, CMV-lucif
31	552.8	31.3	5620	22	AA32866	Minimal adenovirus
32	552.8	31.3	5789	20	AA32867	PGL2-Promoter gene
33	552.8	31.3	5791	20	AA32868	PGL2(GST-ARE)SV40/
34	552.8	31.3	5793	20	AA32869	PGL2(apo AI-ARE)SV
35	552.8	31.3	5793	20	AA32870	PGL2(apo AI-mARE)S
36	552.8	31.3	5818	21	AA32871	Plasmid pGL2 (apoA
37	552.8	31.3	5819	20	AA32872	PGL2(apo AI-250)lu
38	552.8	31.3	5819	20	AA32873	PGL2(apo AI-250)lu
39	552.8	31.3	5819	21	AA32874	Plasmid pGL2 (apoA
40	552.8	31.3	5938	21	AA32875	PGL2(apo AI-491)lu
41	552.8	31.3	6092	20	AA32876	Plasmid pGL2 (apoA
42	552.8	31.3	6092	21	AA32877	Plasmid pGL2 (apoA
43	552.8	31.3	6170	16	AA32878	Vector p19LUC sens
44	552.8	31.3	6254	16	AA32879	TGF-beta responsiv
45	552.8	31.3	6464	16	AA32880	Vector HBVLuc. Sy

ALIGNMENTS

RESULT 1  
AAF24923/C  
ID AAF24923 standard; DNA; 2193 BP.

XX AAF24923;

XX AC

DT 30-APR-2001 (first entry)

DE Complement of Fc epsilon receptor-luciferase fusion protein DNA.

XX Fc epsilon receptor; Fc epsilon receptor; immunoglobulin E; IgE; atopic disease;

KW luminescence inducing protein; allergy; hyper IgE syndrome; chimera;

KW internal parasite infection; B cell neoplasia; luciferase; ss.

XX Chimeric - Homo sapiens.

OS Chimeric - Photuris pennsylvanica.

XX WO200104310-A1.

PD 18-JAN-2001.

PF 13-JUL-2000; 2000MO-US19070.

XX 13-JUL-1999; 99US-0143612.

PR 02-MAR-2000; 2000US-0186412.

XX (HESK-) HESKA CORP.

PA (PROM-) PROMEGA CORP.

XX Weber ER, Wood KV, Hall MP;

PI WPI; 2001-103082/11.









```
Db 858 CGAGTGTCTTAATGCACACGTTTGAAGAAAAACATATTTTACAATCATTTACAAGATTAT 799
Qy 857 aaaaatccccactattgttagtgccctccagttatgggtgttttctgctaaaaagcccat 916
Db 798 AAAGTGGAAAGTACTTACTGTACCAACATTAATGGCATTTTTCACAAAAGTCGGTTA 739
Qy 917 gtcgatcaatacagattatcagacttaacggaagtgtctactctggaggagctctcttagga 976
Db 738 GTTGAAGAGTACGATTTATCGCATTAAGAAATATGTCATCTGGTGGCGCACCTTTATCA 679
Qy 977 aaagatgtcgcagagcagtagcaagaaggttgaattacacctgaatcatacaagaatat 1036
Db 678 AAAGAAATTTGGGAGATGGTGAAGAAACGGTTTAAATTAACATTTTGTGAGCAAGGGTAT 619
Qy 1037 ggattaaactgaactgtctgcgttaagtattaccctcctcataatgtctgtgaaacacaggt 1096
Db 618 GGATTAACAGAAACACCTTCGGCTGTTTAAATTAACCGGACACTGACGTACAGCCGGGA 559
Qy 1097 tcaactggaagcccttgccatacatataaagctaaagtttttagataaagctactctgggaag 1156
Db 558 TCAACTGTAAATATAGTACCAATTTCCAGCTGTAAAGTTGTGGATCTTACAACAGCAAAA 499
Qy 1157 ggcgtagaccagagagaagagcgaataatgctttcaaatgaaagtgaattatgaagaa 1216
Db 498 ATTTTGGGCCCAATGAACACTGGAGATTGTATTTAAAGGCGACATGATATAATGAAGT 439
Qy 1217 tattacaacatccggaagcaactattgatactattgacaaagatgggtgcttcattct 1276
Db 438 TATTATAATGAAGAGCTACTTAAGCAATTTATTAACAAGACGGATGGTGGCTCT 379
Qy 1277 ggagattatggtatatacgaagaagatggaaattctttatagttgatcgtatgaaaga 1336
Db 378 GGTGATATTGCTTATTATGACAATGATGCCATTTTATATTGTGGACAGGCTGAATCA 319
Qy 1337 ctattaaatacaagggatatacagttcgcctgctgaactggaagaaatctgctttacaa 1396
Db 318 TTAATTAATATAAAGGTATACAGTTGACGTTGCCTGCTGCTGAAATTTGAGGGAATACTCTTACAA 259
Qy 1397 catcceaagtattgctgaggggtgttactgaggttcctcggaagaaatttggtggacaatta 1456
Db 258 CACCGCTATATTGTTGATGCGCGCTTACTGTGTATCCGGATGAAGCCGCGGCGAGCTT 199
Qy 1457 cctgctgtctgtgtgttagaactctggcaagcgtgactgactgaaagaaagttcaagat 1516
Db 198 CCAGCTGCAGGTGTGTGTAGTACAGACTGGAATAATCTTAAACGAACAAATCTACAAAAT 139
Qy 1517 ttattgcagcacaagtcactccaagaagcatcttcgagcggtgtcgtatttttagac 1576
Db 138 TTGTTTCCAGTCAAGTTTCAACAGCCCAATGGCTACGTTGGTGGGGTGAAATTTTGGAT 79
Qy 1577 agtattccgaagggccctacttggaataactcatcagaagaggagctccgagaaatatgtgc 1636
Db 78 GAAATTCACAAAGGATCAACTGGAANAATTTGACAGAAAGTGTAAAGCAAAATGTTGAA 19
Qy 1637 cagcagacacacaaa 1650
Db 18 AAACACAAATCTAA 5
```

## RESULT 4

AAF24919

ID AAF24919 standard; cDNA; 1638 BP.

XX

XX

AC AAF24919;

XX

DT 30-APR-2001 (first entry)

XX

KW Nucleotide sequence of a firefly luciferase protein.

XX

KW Fc epsilon receptor; Fc epsilonR; immunoglobulin E; IgE; atopic disease;

KW luminescence inducing protein; allergy; hyper IgE syndrome;

KW internal parasite infection; B cell neoplasia; luciferase; ss.

XX Photuris pennsylvanica.

XX

FH Key Location/Qualifiers

CDS 1..1638

FT /\*tag= a

FT /product= "luciferase"

XX

PN WO200104310-A1.

XX

PD 18-JAN-2001.

XX

XX 13-JUL-2000; 2000WO-US19070.

PF

PR 13-JUL-1999; 99US-0143612.

XX

XX 02-MAR-2000; 2000US-0186412.

XX

PA (HESK-) HESKA CORP.

XX

PA (PROM-) PROMEGA CORP.

XX

PI Weber ER, Wood KV, Hall MP;

XX

DR WPI: 2001-103082/11.

XX

XX P-PSDB; AAB31589.

XX

PT A fusion protein, comprising an Fc epsilon receptor domain and a

XX

PT luminescence inducing protein domain that induces a LP substrate to

XX

PT emit light when contacted with the LP domain, useful for detecting

XX

PT immunoglobulin (Ig) E

XX

XX Claim 12; Page 68-70; 105pp; English.

XX

XX The present sequence encodes a firefly luciferase protein. The protein

XX

CC is used to produce fusion proteins of the invention. The specification

XX

CC describes fusion proteins which comprise a human Fc epsilon receptor

XX

CC (Fc epsilonR) alpha-chain domain and a luminescence inducing protein

XX

CC domain that induces a substrate to emit light when contacted with the

XX

CC immunoglobulin (Ig) E. The fusion protein may be used to detect IgE.

XX

CC It may also be used to identify a compound capable of inhibiting

XX

CC Fc epsilonR protein activity. IgE antibody production is indicative of

XX

CC diseases such as allergies, atopic disease, hyper IgE syndrome, internal

XX

CC parasite infections and B cell neoplasia. Detection of IgE production in

XX

CC an animal following therapy is indicative of the efficacy of the

XX

CC treatment, for example when using treatments intended to disrupt IgE

XX

CC production.

XX

SQ Sequence 1638 BP; 559 A; 254 C; 319 G; 506 T; 0 other;

Query Match 32.5%; Score 573.8; DB 22; Length 1638;

Best Local Similarity 60.0%; Pred. No. 3.6e-102;

Matches 978; Conservative 0; Mismatches 647; Indels 6; Gaps 1;

Qy 26 atggaagaagaacattagcgatggagagcgctcctcgatgatatagtcacctggtcg 85

Db 1 atggaagataaaatttttatatgacgtgacgtgaacctttcatcctctgctgagtgagcg 60

Qy 86 gcaggacaacaattataccattgtataataattgcatcttttccctgaa-----gca 139

Db 61 gctggagacagatgttttaacgattatctcggtgacagatatttcaggatgcatcgca 120

Qy 140 ataactgctgcacaaaatgaagtaataatcatatcatatgcataaataatttgaaaccagctgc 199

Db 121 ttgacaaatgctcacaacaaagaaatgttttatatgaagaatttttaaaattgctgt 180

Qy 200 cgttagctgttagtatagaaacaatagctggttgaatgaacaaatgttgggtgtatgc 259

Db 181 cgttagcggaagtttttaaaagtatgattgataacaaacacacataagcgtgtgt 240

Qy 260 agtgaacaaatataaaacttttttaactcctgctcctgctgttataacttaagaataacca 319

Db 241 agtgaagaatggttgcgaatttttctccttcttaattgcatcatctgttgcgaataatt 300



Query Match		32.4%	Score 571.6;	DB 20;	Length 1642;
Best Local Similarity		59.9%;	Pred. No. 9.5e-102;		
Matches 977;		Conservative 0;	Mismatches 649;	Indels 6;	Gaps 1;
Qy	25	aatggaagaagaacacattagcagatgagagcgtccctcgtgatatagtccatcctggctc	84		
Db	7	aatggaagataaaaaattatttataatggacctgaacctatttatcccttggctgtagtggac	66		
Qy	85	ggcaggacaacaattataccaactcattgtataaaattgcatcttttctctg-----aagc	138		
Db	67	ggctgggaacacagatgttttacgcattatctcgtttatgcagatatatttcaggatgcatagc	126		
Qy	139	aataatcgtatgctcatcaaaaatgaagtataatcatatgctcaaatatttgaacacagctg	198		
Db	127	attgacaaatgctcatcaaaagaagaaatgttttatagaagagtttttaaaattgtcgtg	186		
Qy	199	ccgcttagctgttagtatagaacaatatggcttgaatgaataaaacaacatgttgggtgtgatg	258		
Db	187	tcgcttagcggaaagtttcaaaagtatggataaaacaaacagacacaaatagcgggtg	246		
Qy	259	cagtgaacaacaataaaacttttaactcctcctgtccttgccttatacttaggaataacc	318		
Db	247	tagcgaataatggttgcgaatttttcccttccctttaaattgcacattgtatcttggaaataat	306		
Qy	319	agtagcaacatcaaatgatgatacacagatggagagttaaactggctcatcttgaatatatc	378		
Db	307	tgccgcctgttagtgataaaatacatgaacgtgaattataacacacagctctgtgtattgt	366		
Qy	379	aaaaccaactatcattgttagtctcaagaagaagcactccgctttattctgagagtagacgca	438		
Db	367	aaaaccgcataaattttgttccagaataacttttcaaaaagtactgtaagttaaaatc	426		
Qy	439	aaatctaagtttcatataaaagtcgttagttatcgtatgacgtgtacgacattaatggcgt	498		
Db	427	taaatataaatagtagaacactattattatatagacttaaatgaagacttaggaggtta	486		
Qy	499	tgaatgcgtatctacccttgttgcagcttatactgacacacacacacacacacacacacac	558		
Db	487	tcaatgctcacaacactttatttccaaaatccgatatataatcttgacgtaaaaaaatt	546		
Qy	559	tacaccaaaagattttgatcccttgaataaatcgataattatgtatcatcatctggaac	618		
Db	547	taacacaaatctttaaatacgaacagatcaggttgcgttggtgaatgttttctctggtac	606		
Qy	619	aactggattgctcaaggttgtagtagcgcacatagaagtcataactataagaattcgttca	678		
Db	607	aactgggtttcgaaggagtcgctcaactcacaagaataattgttgcacgattttctca	666		
Qy	679	tagcaggatcccaatttatggcactcgtacggttccacaaacacacacacacacacacacac	738		
Db	667	tgcgaagatcctacttttggtaacgcaatttaattccacgacagcaatttttaacgctaat	726		
Qy	739	accgttccatcatcgccttggaaatgtttactacattattcttactttagtagtagacttaa	798		
Db	727	accttccacacatgggtttggtagatgaccacattaggatacttacttcttggattccg	786		
Qy	799	gggttgtaattgtgaagaataattgaggggcgacacttttcttaaaaccacacagaattacaa	858		
Db	787	agttgctctaataatgcacagctttgaagaaaaactttctacaactatcacagaattataa	846		
Qy	859	aatccccactatttagtggccctccagttatggtgttttggcttaaaagcccaattagt	918		
Db	847	agtggaagatcttacttgtaccacacattaatggcatttttgcanaaagtgcattagt	906		
Qy	919	cgaataacagatttatcagcttaacggaagtgtcctactggaggagctctcttaggaaa	978		
Db	907	tgaagaatcacgatttatcgcactttaaagaaataatgcatctgtgtggcgacctttatacaa	966		
Qy	979	agatgtcgcagacagtagcaaaagaggttgaataatccttgaatcacaacagatagg	1038		
Db	967	agaaattggggagatggttgaagaaacaggttataataactttgtccaggcaaggttatg	1026		
Qy	1039	attaactgaacttgcgtgcgtgtaatgattaccctccataatgctgtgaaacacagattc	1098		

Db	1027	attacagaacccacttgcgtgttttaattacacccgacacactgacgtcagaccggggtc	1086
Qy	1099	aactgggaagacccttgcatacatattaaagctaaagtttttatagataacgcgtactgtggaaagc	1158
Db	1087	aactggtaaaatagtaaccttaccgctgttaaagtgtcgtatcgtacacacaggaataat	1146
Qy	1159	gctaggaccagaagaagagcgcaaaatagcttttcaaaagtgaataattgataaagagata	1218
Db	1147	tttggggccaaatgaacctggagaattgtattttaaagcgcacatgataatgaagaatta	1206
Qy	1219	ttacaacaatcggaagcaactattgatactattgacaaaagatggttgcctcattctcgg	1278
Db	1207	ttataataatgaagaagctactaaagcaattattacaagaacggatggttgcgctctcg	1266
Qy	1279	agatattggatattcacgcaggaagtgaatttctttatagttatgctgattgaaagaact	1338
Db	1267	tgatatgtcttattatgacaatgatggccatttttattatgttgacagggctgaagtcatt	1326
Qy	1339	tattaaatacagaaggtatcaggttgcctgctgaaacgctgaaataatcgtcttttacaaca	1398
Db	1327	aattaaataaaagttatcaggttgcacctgctgaatttgagggaatactactttacaaca	1386
Qy	1399	tccaagtattcgtatgcgggtgttactgaggttccggacgaatttggcgacaattacc	1458
Db	1387	tccgtatattgtatgcgggttactgtgtatccggatgaagccggtgcggagctcc	1446
Qy	1459	tgctgcttgttgtttagaattctgcaagcgtctgactgaaaggaagtccaagattt	1518
Db	1447	agctgcaggttgttagtacagactggaaaatatataacacgaacaaatcgtacaaaaattt	1506
Qy	1519	tattgcagcacaaagtcactccacaaacatcttcgagcggtgtcgtattttgtagacag	1578
Db	1507	tgttccagtcagtttcaacagccaaatggctacgtgtgtgggtgaaaatttttgatga	1566
Qy	1579	tattccgaaagccctactgaaac	1638
Db	1567	aattcccaaaaggaacaaactggaaaatttgacagaaaagtgttaagacaaaattgttga	1626
Qy	1639	gcgagcaccaaa	1650
Db	1627	acacaaatctaa	1638
RESULT 6			
AAS00896			
ID	AAS00896	standard;	cDNA; 1642 BP.
XX	AC	AAS00896;	
XX	DT	04-JUL-2001	(first entry)
XX	DE	Beetle cDNA encoding luciferase mutant LucPpe2 [T249M].	
XX	KW	Beetle; Luciferase; thermostability; recursive mutagenesis;	
XX	OS	ATP detection; luminescence; mutant; LucPpe2; T249M; ss.	
XX	OS	Photuris pennsylvanica.	
XX	OS	Synthetic.	
XX	Key	Location/Qualifiers	
FT	CDS	8..1642	
FT	FT	/*tag= a	
FT	FT	/product= "LucPpe2 [T249M]"	
FT	FT	/partial	
FT	FT	/note= "No stop codon"	
FT	FT	/transl_except= (pos:1202..1204,aa:Gly)	
FT	mutation	replace (753,C)	
XX	XX	WO200120002-A1.	
XX	PD	22-MAR-2001.	



QY 1639 gcagagacacaaa 1650  
 Db 1627 acacaaatctaa 1638

# RESULT 7

ID AAX32847 standard; DNA; 1639 BP.

XX AC AAX32847;

XX 28-JUN-1999 (first entry)

XX Mutant luciferase encoding DNA.

XX Luciferase; enzyme; thermostability; beetle luciferase; ATP assay;  
 KW luminescent; genetic reporter; enzyme immobilisation; mutant; ss.

XX Photuris pennsylvanica.

XX Synthetic.

XX WO9914336-A2.

XX 25-MAR-1999.

XX 18-SEP-1998; 98WO-US19494.

XX 19-SEP-1997; 97US-0059379.

XX (PROM-) PROMEGA CORP.

XX PA Hall MP, Wood KV;

XX WPI; 1999-229538/19.

XX P-PSDB; AAY06854.

XX New mutant luciferase enzymes with increased stability

XX Claim 16; Page 81; 167pp; English.

XX The invention relates to a mutant luciferase enzyme that have increased thermostability. The mutant luciferase is obtained by mutating a polynucleotide sequence encoding a beetle luciferase and selecting expression products having increased thermostability. The new luciferases can be used in ATP assays, as luminescent labels for nucleic acids, proteins or other macromolecules, as genetic reporters, in enzyme immobilisation, as hybrid proteins, in high temperature reactors, and in luminescent solution. The improved thermostability allows storage of luciferases without altering its activity, and improves reproducibility and accuracy of assays using the new luciferases. The novel luciferases can also have improved brightness of luminescence, substrate utilisation and luminescence signal. Sequences AAX32843-856 represent DNA molecules encoding the mutant luciferase enzymes.

XX Sequence 1639 BP; 559 A; 254 C; 324 G; 502 T; 0 other;

Query Match 32.3%; Score 570; DB 20; Length 1639;  
 Best Local Similarity 60.0%; Pred. No. 1.9e-101;  
 Matches 970; Conservative 0; Mismatches 640; Indels 6; Gaps 1;

QY 25 aatggaagaacaaacattagcgatggagagcgctcctcgatagatagtcctcctggctc 84  
 Db 7 aatggaagataaaaattatttatggacctgaaccatttatcccttggctgagtgac 66  
 QY 85 ggcaggacacaaattatcaccaattgataaaaattgcatctttctctg-----aagc 138  
 Db 67 ggcaggacacagatgtttgacgcatattctcgttatgcagataattcagagtgcatgac 126  
 QY 139 aataatcgatgctacatacaaatgaagtaataatcatatgctcaaatatttgaaccagctg 198  
 Db 127 attgacaaatgctacatacaaaagaaattgttttatatgaagatttttaaaattgctg 186

QY 199 ccgottagctgttagtatagaaacaatatggttgaaagaaacaacaatatgttgggtgtatg 258  
 Db 187 tcgttagcggaagtgttaaaagtagtgattaaacaaacacacacaaatagcggtgtg 246  
 QY 259 cagtgaacaaacataaaactttttaaactcctgtcctgtcctgtcttatacttaggaatacc 318  
 Db 247 tagcgaaaatggttgcgaatttttccctccataaattgcatactgttatcttggataat 306  
 QY 319 agtagcaacatcaaatgatgatgtcacacagatggagagtaactggtcatttgaatatatc 378  
 Db 307 tgcgacacctgttagtgataaatacatgaacgtgaattaaatacacacagctcttggatgt 366  
 QY 379 aaacccaactatcattgttttagttaaagaagacaccccgcttatttctgagagtagaca 438  
 Db 367 aaacccacgcataatttttctccagaatacttttcaaaaagtactgaatgtaaatc 426  
 QY 439 aaactaagtttcatataaaaagtcgttagttatcgatagcgttacgacattaaatggcgt 498  
 Db 427 taattaaaaatatgagaactattattattagaaactaaatgaagacttaggaggtta 486  
 QY 499 tgaatgcglatcctaocctttgttgacgcttatctgaccacaccccttgcattcatttgc 558  
 Db 487 tcaatgcctcaacaactttattctcaaaattccgataattaatcttgcgttaaaaaatt 546  
 QY 559 tacacaaaagattttgatcccttgaaaaaatcgcatataattatgtcatcatctcggaac 618  
 Db 547 taacattattctttaaactgagacgacaggttgcgttgcgttaattgtttctcgtgtac 606  
 QY 619 aactggattgcctaaggtgttagtactgacatagacagcttaactataaagattcgttca 678  
 Db 607 aactgggttttcgaaggagtcactgtaactcaagaataattgttgcacgatttctca 666  
 QY 679 tagcagggatcccatattttagcactcgtcaggtttccacaaacatcaattcttcttagt 738  
 Db 667 tgcaaaagatcctactcttggtaacgcaattaatccacacagacgaattttaacggtta 726  
 QY 739 accgttccatcatgccttttggaaatgtttactacattatcttacttctgttagtaggactaa 798  
 Db 727 accttccaccatggttttggtagatgaccacattagacttacttcttctgttggattccg 786  
 QY 799 ggttgttaattggaagaatttgaggcgcacttttctaaaaaccatacagaatacaaa 858  
 Db 787 agttgttctaatagcacggtttgaaagaaaacattttacaaatcattacaagattataa 846  
 QY 859 aatccccattttagtggccctccagttatggtgttttggctaaagccccattagt 918  
 Db 847 agtggaaagtactttacttacttaccacaaattaatgcaatttttgcacaaagtgcattag 906  
 QY 919 cgatcaatcacgatttatcgagcttaacggaagtgtcactgagagagctcctttaggaaa 978  
 Db 907 tgaaaagcacgatttatcgacacttaaaagaaattgcactctgtggtggcgacacctttatca 966  
 QY 979 agatgtcgcagaagcagtagcaaaaggttgaatttacctggaatcacaaggataggg 1038  
 Db 967 agaaattggggagatggtgaaaaaacggtttaaaattaaacttgcaggcgaagggtatgg 1026  
 QY 1039 attaaactgaaacttgcgtgcgtgtaattgattaccctccataatgctgtgaaacacgggttc 1098  
 Db 1027 attaacagaaacacacttcggctgttttaattacacccgaacaatgacgtcagaccggatc 1086  
 QY 1099 aactggaagacccttgcctacatacattaaagctaaaagtttttagataacgctactcgggaaggc 1158  
 Db 1087 aactggttaaaatagcacctttccacgtgtttaaagtttgcgactcacaacgggaaaaat 1146  
 QY 1159 gctagggaccaggagaagggcgaaatgctttcaaaagtgaattgattgaaaggata 1218  
 Db 1147 ttgggggccaataatgaactggagaattgtattttaaagggcacatgataatgaagggtta 1206  
 QY 1219 ttacaacaatccgggaagcaactattgtactatttgcacaaagatggttggcttcttcctgg 1278  
 Db 1207 ttataataatgaagaagctactaaagcaattatttaacaaagcggatggttgcgtctcgg 1266  
 QY 1279 agatattggattattacgacggaagatggaaaattttctttattagttgcattgaaagaact 1338

Db 1267 tgaattgctcatatgacaatgatggccattttatatattgtgacaggtcgaagtcatt 1326  
Qy 1339 tatataacacaggagatacaggttgccctgctgaaactggaaatcgtctttacaaaca 1398  
Db 1327 aattataataaaggattacaggttgccctgctgaaattgagggaatactcttacaaca 1386  
Qy 1399 tccaagtattgctgatcggtgttacttgagggttcccgacgaatattggtgacaaattacc 1458  
Db 1387 tccgtatatattgtagtcggttacttggtataccggtgaagccggtgagcttcc 1446  
Qy 1459 tgcgtctgttattggttagaatactggcaagcgtgactgaaggaagttcgaattt 1518  
Db 1447 agctgcaggtgtttagtacagactggaaatacttaacacgaacaaatcgtaacaaattt 1506  
Qy 1519 tattgcagcacaagtcactccaaacacatcttcgagcgtgtcgtattgttagacag 1578  
Db 1507 tgtttccagtcagatttcaacagcaaatggctacgtggtggtggaatttttggatga 1566  
Qy 1579 tattccgaagccctacttgaaactcatcagaaaggagctccgagaaatatttg 1634  
Db 1567 aattcccaaggatcaactggaaatttgacagaaagggttaagacaaatgtttg 1622

RESULT 8

ID AAX32844  
XX AAX32844 standard; DNA; 1639 BP.

AC AAX32844;

DT 28-JUN-1999 (first entry)

DE Mutant luciferase encoding DNA.

KW Luciferase; enzyme; thermostability; beetle luciferase; ATP assay;  
KW luminescent; genetic reporter; enzyme immobilisation; mutant; ss.

OS Photuris pennsylvanica.

OS Synthetic.

XX WO9914336-A2.

XX 25-MAR-1999.

XX 18-SEP-1998; 98WO-US19494.

XX 19-SEP-1997; 97US-0059379.

XX (PROM-) PROMEGA CORP.

XX Hall MP, Wood KV;

XX WPI; 1999-229538/19.

XX P-PSDB; AAY06851.

XX New mutant luciferase enzymes with increased stability

XX Claim 16; Page 79; 167pp; English.

XX The invention relates to a mutant luciferase enzyme that have increased  
XX thermostability. The mutant luciferase is obtained by mutating a  
XX polynucleotide sequence encoding a beetle luciferase and selecting  
XX expression products having increased thermostability. The new  
XX luciferases can be used in ATP assays, as luminescent labels for nucleic  
XX acids, proteins or other macromolecules, as genetic reporters, in enzyme  
XX immobilisation, as hybrid proteins, in high temperature reactors, and  
XX in luminescent solution. The improved thermostability allows storage of  
XX luciferases without altering its activity, and improves reproducibility  
XX and accuracy of assays using the new luciferases. The novel luciferases  
XX can also have improved brightness of luminescence, substrate utilisation  
XX and luminescence signal. Sequences AAX32843-856 represent DNA molecules  
XX encoding the mutant luciferase enzymes.

Qy 25 aatggaagaagaacattagcgatggagcgtctcctcgtgatatagtcctatcctggctc 84  
Db 7 aatggaagataaataattttatataggacatgaacccattttatccttggctgagggac 66  
Qy 85 ggcgggacacaaattataccaaatcattgtataaaatttgcatacttttctctg-----aagc 138  
Db 67 ggcgggacacagatgtttacgcattatctcgtattatcagataattcaggatgcatagc 126  
Qy 139 aataatcgatcctcatcacaaatgaagtaatacatcatatctcaaatatttgaaccagctg 198  
Db 127 attgacaatgctcatcacaaagaaaatgtttatatgaaggttgttaaaattgtctg 186  
Qy 199 ccgcttagctgttagatagacaataatggcttgaatgaacaaatgttggtgggtatg 258  
Db 187 tcggttagcggaagttttaaagataggtataaaacacaaacagacacaaatagcgtgtg 246  
Qy 259 cagtgaacacaaatataaaacttttttaactcctgtcctgtctgttatacttaggaatacc 318  
Db 247 tagcgaaatggtttgcaatttttctcctataaattgcatactgtatcttctggaataat 306  
Qy 319 agtagcaacatcaaatgatgtacacagatggagaggttaactgctatttgaatatatc 378  
Db 307 tgcagcactgttagtgataataatcattgaacgtgaattatacacacagctctgttctg 366  
Qy 379 aaacccactatcatgttttagttcacaagaagcactccgcttattcttgagagtcagaca 438  
Db 367 aaacccagcaataatttttctcacaagaatacttttcaaaagctactgaaatgtaaaatc 426  
Qy 439 aaatcctagttcatlaaaaaagtcgtagtattatogatagcatgtacgacattaaaggcgt 498  
Db 427 taaattaaaatagttagaaactattattattatagactaaatgaagacttagagggtta 486  
Qy 499 tgaatcgatctaccctttgtgcacgttatctactgacacacacaccttgcacatttctatt 558  
Db 487 tcaatgcctcaacactttatttccaataatccogataattatcctgggacgtanaaaaaatt 546  
Qy 559 tacacaaaagattttgaccccttgaaaaaatcgcatattatgtcatcatctgggaac 618  
Db 547 taaaccatactcttttaactgagacgatacaggttgcgttggttaatgttttctctcgttac 606  
Qy 619 aactggattcctcaagggtgttagtactgagccatagaaagtctaaactaagaattcgttca 678  
Db 607 aactggtgtttcgaaggagtcagctaaactcacagaataattgttgacgattttctca 666  
Qy 679 tagcagggatcccatattatggcactcgtacggttccacaaacacataatcttctcattagt 738  
Db 667 tgcaaaagatcctactcttttggtaacgcaatttaatccacacgacagcaattttaaocggta 726  
Qy 739 accgttcacatgccttctggaattttactacattatcttacttcttagtagagacttaa 798  
Db 727 acccttccacatggtttctggtatgatgacacacattagatacttacttacttggattccg 786  
Qy 799 ggttgaattgttgagaataatttgaggcgcacttttcttaaaacccatacagaattacaa 858  
Db 787 agtgttctaatagcacacgctttgaagaaaaaactatttacaacatttacaagattataa 846  
Qy 859 aatccccactatttagtggccctccagttatggtgttttttggctaaaaagcccaattagt 918  
Db 847 agtggaaagtcatttactgttaccacacattaatggcatttttttgcacaaagtgcattagt 906  
Qy 919 cgatacaacacattatctcagcttaaacgaaagtgtcactgagagagctcctcttaggaaa 978  
Db 907 tgaagaagcacagatttatcgacttaaaagaataatgcattctggtggcgacattttacaa 966  
Qy 979 agatgtcgcagaagcagtagcaaaagaggttgaatttacctggaatcacaagaagatttgg 1038

XX



Db 967 agaaattgggagatggtgaaaaaacgggttttaataataaactttgtcaggcaagggtatgg 1026  
QY 1039 attaaactgaactgtgcgtgctaaagtattacccttcataatgctgtgaaacaggttc 1098  
Db 1027 attaacagaacaccttcggctgttttaattacacccgaaacaatgacgtcagaccggatc 1086  
QY 1099 aactgggaagaccttgcatacatataaagctaaagtatttagataaacctactggaagcc 1158  
Db 1087 aactggttaaatagtaaccttaccgtgtttaagtgtcgtacctcaacagagaaat 1146  
QY 1159 gctaggaccagggaagagcggaatatgctttcaagtgaatgattgattgaagata 1218  
Db 1147 ttggggccaaatgaactggagaatgtattttaaaggcgacatgataatgaaggtta 1206  
QY 1219 ttacaacaatccggaacactatgatactattgatacgaagatggttgccttcaattcg 1278  
Db 1207 ttataaatgaagaagctactaagaacattattatacaaaagacggatggttgcgctctg 1266  
QY 1279 agatattggtattacagacgaagatgaaatttcttattagttgatcgattgaaagaact 1338  
Db 1267 tgatattgcttattagacaatgatggccatttttattatgtggacaggctgaagtcat 1326  
QY 1339 tattaaatacaagggtatcagggttgcgctgctgtaactggaatactgcttttacaaca 1398  
Db 1327 aattaaataaaggttatcagggttgcacctgtgaaattgagggaatactcttacaaca 1386  
QY 1399 tcaagttatgtgatcggtgttactgaggtcccggaagatatttggtagacaattacc 1458  
Db 1387 tccgtattgttatgcggcttactgtgtataccggtgaaagccggcgagcttcc 1446  
QY 1459 tgcgtctgtgtgttagaatctggaagacgtcactgaaaggaagttcaagattt 1518  
Db 1447 agctgcagggtgtgtgacagactggaaaatatctaaacgaaacaaatcgacaaaattt 1506  
QY 1519 tattcgagcacaagctcactccaaacaaagcatcttcgaggggtgtcgtattttagacag 1578  
Db 1507 tgttccagcaagtttcaacagcagcaaatggtacgtggtgggtgaaattttggatga 1566  
QY 1579 tattccgaagggccctactggaactcaactcaatcagaaaggagctccgagaataatttg 1634  
Db 1567 aattcccaaggatcaactggaataattgacagaaaagtgtaaagaaaattgtttg 1622

RESULT 9  
AAX32843  
ID AAX32843 standard; DNA; 1639 BP.  
XX AC  
XX AC AAX32843;  
XX DT 28-JUN-1999 (first entry)  
XX Mutant luciferase encoding DNA.  
DE Luciferase; enzyme; thermostability; beetle luciferase; ATP assay;  
KW luminescent; genetic reporter; enzyme immobilisation; mutant; ss.  
XX OS Photuris pennsylvanica.  
XX Synthetic.  
XX PN W09914336-A2.  
XX PD 25-MAR-1999.  
XX PF 18-SEP-1998; 98WO-US19494.  
XX PR 19-SEP-1997; 97US-0059379.  
XX PA (PROM-) PROMEGA CORP.  
XX PI Hall MP, Wood KV;  
XX WPI; 1999-229538/19.  
XX P-PSDB; AAY06850.

XX New mutant luciferase enzymes with increased stability  
XX Claim 16; Page 79; 167pp; English.  
XX The invention relates to a mutant luciferase enzyme that have increased  
CC thermostability. The mutant luciferase is obtained by mutating a  
CC polynucleotide sequence encoding a beetle luciferase and selecting  
CC expression products having increased thermostability. The new  
CC luciferases can be used in ATP assays, as luminescent labels for nucleic  
CC acids, proteins or other macromolecules, as genetic reporters, in enzyme  
CC immobilisation, as hybrid proteins, in high temperature reactors, and  
CC in luminescent solution. The improved thermostability allows storage of  
CC luciferases without altering its activity, and improves reproducibility  
CC and accuracy of assays using the new luciferases. The novel luciferases  
CC can also have improved brightness of luminescence, substrate utilisation  
CC and luminescence signal. Sequences AAX32843-856 represent DNA molecules  
CC encoding the mutant luciferase enzymes.  
XX SQ Sequence 1639 BP; 558 A; 255 C; 323 G; 503 T; 0 other;

Query Match 32.1%; Score 566.8; DB 20; Length 1639;  
Best Local Similarity 59.9%; Pred. No. 8.1e-101;  
Matches 968; Conservative 0; Mismatches 642; Indels 6; Gaps 1;

QY 25 aatggaagaagaacatttaggcgtgagagcgtcctcgtgatagtagccactcctggtc 84  
Db 7 aatggaagaataaattttatgtgacctgaaaccttttcccttgcgtgatgggac 66  
QY 85 ggcaggacaacaattatccaatcatgttatataatttgcctcttttctg-----aagc 138  
Db 67 ggcaggacaacagatgttttacgcattatctcgttatgcagatatattccaggatcgatgc 126  
QY 139 aataatcgatcctcacaacaaatgaagtaataatcatatctcacaataatttgaacacagctg 198  
Db 127 attgacaaatcctcacaacaaagaaatgttttatgaagagtttttaaaattgctgtg 186  
QY 199 ccgttagctgttagtataagacaataatggttgaaagaaacaatgttgcgggtgtatg 258  
Db 187 tcgtttagcggaaagttttaaaagtgatggaataaaacaaacagacacaatgagcgtgtg 246  
QY 259 cagtgaacacaataaaacttttttaactcgtcgtcgtctttatctacttaggaataacc 318  
Db 247 tagcgaataatggttgcaatttttccctcctcctaataatgcatcatcgtatccttggaataat 306  
QY 319 agtagcaacatcaaatgatgtatcacacagatggagagtgtaactggtcatttgaatatac 378  
Db 307 tgcagcactgttagtgataaatacattgaacgtgaataataacacagcttctgtattgt 366  
QY 379 aaacccaactatcatgttttagttccaagaaagcactccgcttattctgtgaggtacagca 438  
Db 367 aaacccagcgaataattttgtctccaagatacttttcaaaaagtactgaatgataaatac 426  
QY 439 aaactcaagtttcaataaaaagtcgttagttatgcagatgacgtacgacattaaatgcgt 498  
Db 427 taataataaatatgtagaaactattattattagacttaataatgaagacttaggggtta 486  
QY 499 tgaatgcgtatctaccttgttgcacgttatactagcacaacaccccttgcctcattgcatt 558  
Db 487 tcaatgcctcaacaacttatttctcaaaaattccgataattaatcttgacgtataaaaaatt 546  
QY 559 tacacaaaagattttgatcccttgaaaaaatcgcatttaattatgcatcatcctcgtggaac 618  
Db 547 taacataattctttaaactcgagacgatcagggttgcgttggtgaatgttttctcgtgtac 606  
QY 619 aactggattgcctaagggtgtagtagcagccatagaagtcataaactataaagattcgttca 678  
Db 607 aactgggttttcgaaggagtcagttaactcaacaagaataattgttgcacgatttctct 666  
QY 679 tagcagggatcccaatttatggcactcgtacggttcccaacacatacattcttctcctagt 738  
Db 667 tgcaaaagatcctactcttcttggttaacgcgaatttaataccacgacagcaattttaacgggta 726

```
Qy 739 accgttcacatcgcttctggaattgtttactacatttacttacttcttagtagagacttaa 798
  ||| |||| |||| |||| ||| || || |||| |||| |||| |||| |||
Db 727 accttcacacatggttttggatgagaccacatttggatcttacttctggtctccg 786
  |||| ||| || |||| ||| || || |||| |||| |||| |||| |||
Qy 799 ggttgtaattgtgaagaaatttgaggcgacactttcttaaaaaccatacagaatacaa 858
  |||| ||| || |||| ||| || || |||| |||| |||| |||| |||
Db 787 agtggcttaatgcacacgttttgaagaaaaactatttctacaatcattacaagattataa 846
  |||| ||| || |||| ||| || || |||| |||| |||| |||| |||
Qy 859 aatccccactattgtagtgccctccacgttatgtgttttttggctaaaagcccatagt 918
  || || || || || || || || || || || || || || || || || || || || || ||
Db 847 agtggaaagtacttacttctgaccacatttaattggtcattttctgcaaaaagtcattagt 906
  || || || || || || || || || || || || || || || || || || || || || ||
Qy 919 cgatcaatagatttctagcttaacggaagtgtctactgagagactccttttagaaa 978
  || || || || || || || || || || || || || || || || || || || || || ||
Db 907 tgaagtagcatttctgcactcaagaatttgcactcgtggtgcgacacttctatcaa 966
  || || || || || || || || || || || || || || || || || || || || || ||
Qy 979 agatgcgcagaagcagtagcaagaggttgaaatttaccctggaatcatacaagatatgg 1038
  || || || || || || || || || || || || || || || || || || || || || ||
Db 967 agaaattgggagatggtgaaaaaacggttttaataataaactttgcaaggcaagggtatgg 1026
  || || || || || || || || || || || || || || || || || || || || || ||
Qy 1039 attaaactgaaactgtcgcgtctgtaagattaccctccatataatcgttgaaaaacaggttc 1098
  |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||
Db 1027 attaacgaaaacactcggctgttttaattacaccgaaacaatgacgtcagacgggatc 1086
  |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||
Qy 1099 aactggaagacccttgcacatacattaaagctcaaggttttagataacgcactactggaaggc 1158
  |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||
Db 1087 aactggtaaaatagatcacttccgcgtgttaaagttgtcgtactcctcaacaggaaaaat 1146
  |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||
Qy 1159 gctaggaccgggaaagagagcgaataatgctttcaagtgaaatgattatgaagaagata 1218
  |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||
Db 1147 ttggggccaaatgaactggagaatgtgtattttaaggcgacatgataatgaagggtta 1206
  |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||
Qy 1219 ttacaacaatccggaagcaactatgtactactatgacaaagatggttgcttctattcgg 1278
  || || || || || || || || || || || || || || || || || || || || || ||
Db 1207 ttataataatgaagaagctactaaagcaattatttaacaagacggtggtgctcctgg 1266
  || || || || || || || || || || || || || || || || || || || || || ||
Qy 1279 agatattggattatcagcagaagtggaatttctttagttgattgattgaaagaact 1338
  |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||
Db 1267 tgatatgtctattatgacaatgatggccatttttatattgttgacagggctgaagtcatt 1326
  |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||
Qy 1339 tattaatatacagaaggtatcaggttgcgcctgctgaaactggaataatctgctttacaaca 1398
  |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||
Db 1327 aattaaataaagtttatcaggttgccctgctgaaattgagggaaataactcttacaaca 1386
  |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||
Qy 1399 tccaaagtattgctacgggtgttactggaagttccggaagaaatttggtgacaattacc 1458
  || || || || || || || || || || || || || || || || || || || || || ||
Db 1387 tccgtatattgtgatccggcgctactggtataccggatgaagccggcgagcttcc 1446
  || || || || || || || || || || || || || || || || || || || || || ||
Qy 1459 tgctgcttgtgtgttagaatactggcagacgcgtgactgaaaggaagttcaagattt 1518
  |||| |||| || || || || || || || || || || || || || || || || || ||
Db 1447 agctgcagggtgtgttagtacagactggaaaatatctaaacgaacaaatcgtaaaaaattt 1506
  |||| |||| || || || || || || || || || || || || || || || || || ||
Qy 1519 tatgagacaaactcactccaaacaaagcatcttcggaggcggtgcgtattgttagacag 1578
  || || || || || || || || || || || || || || || || || || || || || ||
Db 1507 tgtttccagtcagtttcaacagcgaatggctacgtggtgggtgaaatttttggatga 1566
  || || || || || || || || || || || || || || || || || || || || || ||
Qy 1579 tattccaaaggccctactggaacactcacaagaaagagctccgagaaatatttg 1634
  |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||
Db 1567 aattcccaaggatcaactggaaaaattgacagagaagtgcttaagacaaaatggtttg 1622
  |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||
```

## RESULT 10

AA32845

ID AA32845 standard; DNA; 1639 BP.

XX AC AA32845;

XX AC AA32845;

DT 28-JUN-1999 (first entry)

XX Mutant luciferase encoding DNA.

XX Luciferase; enzyme; thermostability; beetle luciferase; ATP assay;

```
KW luminescent; genetic reporter; enzyme immobilisation; mutant; ss.
XX Photuris pennsylvanica.
OS Synthetic.
XX WO9914336-A2.
XX 25-MAR-1999.
XX 18-SEP-1998; 98WO-US19494.
XX 19-SEP-1997; 97US-0059379.
XX (PROM-) PROMEGA CORP.
XX Hall MP, Wood KV;
XX WPI; 1999-229538/19.
XX P-PSDB; AAY06852.
XX New mutant luciferase enzymes with increased stability
XX claim 16; Page 80; 167pp; English.
XX The invention relates to a mutant luciferase enzyme that have increased
  thermostability. The mutant luciferase is obtained by mutating a
  polynucleotide sequence encoding a beetle luciferase and selecting a
  expression products having increased thermostability. The new
  luciferases can be used in ATP assays, as luminescent labels for nucleic
  acids, proteins or other macromolecules, as genetic reporters, in enzyme
  immobilisation, as hybrid proteins, in high temperature reactors, and
  in luminescent solution. The improved thermostability allows storage of
  luciferases without altering its activity, and improves reproducibility
  and accuracy of assays using the new luciferases. The novel luciferases
  can also have improved brightness of luminescence, substrate utilisation
  and luminescence signal. Sequences AAX32843-856 represent DNA molecules
  encoding the mutant luciferase enzymes.
XX Sequence 1639 BP; 557 A; 255 C; 323 G; 504 T; 0 other;
```

Query Match 32.1%; Score 566.8; DB 20; Length 1639;  
Best Local Similarity 59.9%; Pred. No. 8.1e-101;  
Matches 968; Conservative 0; Mismatches 642; Indels 6; Gaps 1;

```
Qy 25 aatggaagaagaacacattaggaatggagagcgctccctgctgatatagtcacatcctggctc 84
  |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||
Db 7 aatggaagataaaaataatttttatatggacctgaaccattttatcccttggctgattgggac 66
  |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||
Qy 85 ggcaggacaacaattataccacatctatgataaaatttgcactcttttctctg-----aagc 138
  |||| |||| || || || || || || || || || || || || || || || || || ||
Db 67 ggcaggagaacagatgttttacgcatctatctcgttatgcagatatcttcaggatgcatagc 126
  |||| |||| || || || || || || || || || || || || || || || || || ||
Qy 139 aataatcgatgctcacaacaaatgaagtaataatcatatagctcaaaatatttgaaacacagctg 198
  || || || || || || || || || || || || || || || || || || || || || ||
Db 127 attgacaatgctcacaacaaagaaatgttttatagaagagtttttaaaattgtctgtg 186
  || || || || || || || || || || || || || || || || || || || || || ||
Qy 199 ccggttagctgttagtatagacaataatggttgaaagaaaaacaaatgttgggtgtgtatg 258
  || || || || || || || || || || || || || || || || || || || || || ||
Db 187 tcggttagcggaaggttttaaaagtgatggattaaacaaacacacacaaatagcgggtgtg 246
  || || || || || || || || || || || || || || || || || || || || || ||
Qy 259 cagtgaacaaacataaaacttttttaactcctgtccttgcgttttatacttaggaataacc 318
  || || || || || || || || || || || || || || || || || || || || || ||
Db 247 tagcgaataatggttgcgaattttccctcctatataatgcacatctgtatcttggataaat 306
  || || || || || || || || || || || || || || || || || || || || || ||
Qy 319 agtagcaacatcaaatgatatgtacacagatggagagtgtaactggtcatttgaatatatc 378
  || || || || || || || || || || || || || || || || || || || || || ||
Db 307 tgcagccctgttagtataaatacattggaacgtgaattatatacacacagcttgggtattgt 366
  || || || || || || || || || || || || || || || || || || || || || ||
Qy 379 aaaaccaactatcattgttagttcaagaagaagcactcccgcttattctggaggtacagca 438
  |||| |||| || || || || || || || || || || || || || || || || || ||
Db 367 aaaccacgcataaatttttgcgtccagaaatcacttttcaaaaaagtaactgaatgtaaaatc 426
  |||| |||| || || || || || || || || || || || || || || || || || ||
```

Qy	439	aaatctaagt	tttcattaaa	aaagtcgt	agtattcatg	atagcatg	tacgacattaa	tgccgt	498
Db	427	taaattaaa	atctgtaga	aaactatt	tattatattag	acttaaat	gaagacttag	gggtta	486
Qy	499	tgaatgcgt	atctac	cttctgtg	cagcttata	ctagtcac	acacaccttg	tcatctg	558
Db	487	tcaatgcct	caacactt	tatttc	caaaatccg	atatataat	cttgacgt	taaaaatt	546
Qy	559	tacacaa	agagattg	atccccct	gaaaaatcg	catataat	tatgtcat	catctgg	618
Db	547	taaacca	tattctt	taacog	agacatc	agggtgcgt	tggtaatg	ttttctctg	606
Qy	619	aactggat	gcctaa	gggtgt	agtactg	agccata	agaagtc	taactata	678
Db	607	aactgg	tttcga	aggggtc	atgcta	actcac	agaatat	tggtgac	666
Qy	679	tagcagg	atccccatt	atagga	ctcgtag	gtgtcc	acaaacat	caattctt	738
Db	667	tgcaaa	agaccc	ctacttt	ggtaac	gcaatt	atccacg	acgcaatt	726
Qy	739	accgttc	catcatg	ctctt	ggaa	gttttact	acattact	tactt	798
Db	727	acctttcc	accatg	gtttt	ggtagat	gaccac	attagata	cttact	786
Qy	799	ggttg	taattgt	gaaga	aaattt	gagggcg	acacttt	ttt	858
Db	787	agttgt	ctcaatg	cacacg	tttga	agaaa	aaactatt	tt	846
Qy	859	aatcccc	actattg	tagtgg	cccctcc	agttat	gtgtttt	gct	918
Db	847	agtgg	aaagt	acttact	tg	ccccac	attag	tg	906
Qy	919	cgatca	atcag	attat	cagag	cttaac	ggaagt	gtct	978
Db	907	tga	aaagtac	gattat	cgcact	taaaa	gaattgc	atct	966
Qy	979	agatg	tcgc	agaag	cagtag	caaga	aggttg	aaatt	1038
Db	967	agaa	attggg	agatg	gtg	aaaaa	acggtt	taaat	1026
Qy	1039	attaa	ctgaa	actgtg	tcg	ctgtg	taagtatt	acc	1098
Db	1027	attaa	cgaa	acac	ctcg	ctgtgt	ttaatt	ac	1086
Qy	1099	aactg	gaag	accc	ttgc	ctacat	tataa	gctaaa	1158
Db	1087	aactg	gtaaa	aatag	tacc	attt	ccag	ctgt	1146
Qy	1159	gctagg	accgg	gaag	agcg	gaata	tgc	ttt	1218
Db	1147	tttggg	ccaaa	tga	actt	ggga	attgt	tatt	1206
Qy	1219	ttaca	caat	tcgga	agca	actatt	gatata	ct	1278
Db	1207	ttata	ata	tga	aga	agctact	aaag	caatt	1266
Qy	1279	agat	att	ggatat	tcag	cgaga	gtgga	attt	1338
Db	1267	tgat	att	tgct	ctata	tga	caattg	agcc	1326
Qy	1339	tataa	ata	caag	ggata	tcag	gttgc	gc	1398
Db	1327	aatt	aaaata	ta	agg	ttat	tcag	gttgc	1386
Qy	1399	tcaa	agtatt	gctg	atgc	gggtgt	tact	ggag	1458
Db	1387	tc	cg	tatt	gtt	gatgc	cg	cg	1446
Qy	1459	tg	cg	ctgt	gtt	gtt	taga	attc	1518
Db	1447	ag	ctgc	agg	ggt	gt	gt	ag	1506
Qy	1519	tatt	gc	ag	ca	ag	ctc	ac	1578

Db	1507	tggttcacagtcatttcaacagccaaatggctacgtggtgggtgaaatttttgatga	1506
QY	1579	tattccgaaagccctactgtgaaacctcatcagaaaggactccgagaaatatitg	1634
Db	1567	aattcccaaggatcaactggaaaattgacagaaaagtgttaagacaaatgtttg	1622
RESULT 11			
AAS00889			
ID	AAS00889 standard; cDNA; 1639 BP.		
XX	AC	AAS00889;	
XX	DT	04-JUL-2001 (first entry)	
XX	DE	Beetle cDNA encoding luciferase mutant Luc49-4G11.	
XX	KW	Beetle; Luciferase; thermostability; recursive mutagenesis;	
XX	KW	ATP detection; luminescence; mutant; Luc49-4G11; ss.	
XX	OS	Photuris pennsylvanica.	
XX	OS	Synthetic.	
XX	Key	Location/Qualifiers	
FT	CDS	8..1639	
FT	FT	/*tag= a	
FT	FT	/product= "Luc49-4G11"	
FT	FT	/partial	
FT	FT	/note= "No stop codon"	
FT	FT	transl_except= (pos:149..154,aa:Glu>Asn)	
FT	FT	transl_except= (pos:170..172,aa:Phe)	
FT	mutation	replace (86,T)	
FT	FT	/*tag= b	
FT	mutation	replace (150..153,AAA)	
FT	FT	/*tag= c	
FT	mutation	replace (172,T)	
FT	FT	/*tag= d	
FT	mutation	replace (278,T)	
FT	FT	/*tag= e	
FT	mutation	replace (388,T)	
FT	FT	/*tag= f	
FT	mutation	replace (554,A)	
FT	FT	/*tag= g	
FT	mutation	replace (668..670,TGC)	
FT	FT	/*tag= h	
FT	mutation	replace (792,C)	
FT	FT	/*tag= i	
FT	mutation	replace (1064,G)	
FT	FT	/*tag= j	
FT	mutation	replace (1068,C)	
FT	FT	/*tag= k	
FT	mutation	replace (1202,A)	
FT	FT	/*tag= l	
FT	mutation	replace (1632..1635,AATC)	
FT	FT	/*tag= m	
FT	mutation	replace (1637..1639,ANGCTG)	
FT	FT	/*tag= n	
XX	XX	WO200120002-A1.	
PN	XX	22-MAR-2001.	
PD	XX	22-DEC-1999; 99WO-US30925.	
XX	XX	15-SEP-1999; 99US-0396154.	
PF	XX	(PROM-) PROMEGA CORP.	
PR	XX	Wood KV, Hall MP, Gruber M;	
PA	XX	WPI; 2001-244809/25.	
PI	XX	P-PSDB; AAU01223.	

XX Mutant luciferases from Photuris pennsylvanica and Pyrophorus  
PT plagiophthalmus having increased thermostability, increased resistance  
PT to inhibitors and/or enhanced enzymological properties, useful for  
XX detecting ATP -  
PS  
XX Disclosure; Fig 26; 260pp; English.  
XX  
CC The sequence encodes a Beetle mutant luciferase, Luc49-4G11. Mutant  
CC luciferases from Photuris pennsylvanica and Pyrophorus plagiophthalmus,  
CC are created by recursive mutagenesis to have one or more desired  
CC properties, e.g. increased thermostability, increased resistance to  
CC inhibitors and/or enhanced enzymological properties relative to a  
CC wild-type enzyme. The mutant luciferases are useful for detecting ATP,  
CC for labeling a molecule, as genetic reporters, for immobilization onto a  
CC solid surface, to produce hybrid proteins, for high temperature  
CC reactions, or for creating luminescent solutions. The luciferase enzymes  
CC have increased thermostability which allows storage of the enzyme without  
CC altering its activity and improves reproducibility and accuracy of  
CC assays.  
XX  
XX Sequence 1639 BP; 555 A; 258 C; 325 G; 501 T; 0 other;

Query Match 32.1%; Score 566.8; DB 22; Length 1639;  
Best Local Similarity 59.9%; Pred. No. 8.1e-101;  
Matches 968; Conservative 0; Mismatches 642; Indels 6; Gaps 1;

Qy 25 aatggaagaagaacacattagggatggagagcgtcctcgtgatagtagccatcctggctc 84  
Db 7 aatggaagataaaatatatttatatggacctgaacattttatcccttggctgtagggac 66  
Qy 85 ggcaggacacacattatcaccaatcattgtataaatttgcattcttctcgt-----aagc 138  
Db 67 ggcaggacacacattatcaccaatcattgtataaatttgcattcttctcgt-----aagc 126  
Qy 139 aataacatcgtctcatatacaaatgaagtaataatcatatcattgctcaaatatttgaacacagctg 198  
Db 127 attgcaaatcgtcatatacaaaagccctgttttatgatgagagtggtttcaaaattgctg 186  
Qy 199 ccgttagctgttagtagacacaaatagcgttggaatgaacaaatgtgtggtgtagtg 258  
Db 187 cgtttagcgtgaagtttttaaaagtagtggtattaaacaaacagacacaaatagcgtg 246  
Qy 259 cagtgaaacacataaaacttttttaaaccctcctcgtcgtcttcttacttaggaataacc 318  
Db 247 tagcgaataatggttggcaattttctccatataaattgcatcattgtatcttggataat 306  
Qy 319 agtagcaacataaatgatagatgacacagatggagagtgtaactggtcatttgaatatac 378  
Db 307 tgcagcacctgttagtgataaatacatatgaacgtgaatttaataacacagctcttggatgt 366  
Qy 379 aaacacacactatcattgttagttaaagaagacacccctccttattcttgagagtagaca 438  
Db 367 aaacacacacataatttttgcacgaataacttttcaaaagtagtgaatgtaaaatc 426  
Qy 439 aaatttagtcttcataaaagtcgttagttatcgtatgacatgacacataatgaatgcgt 498  
Db 427 taataataaatatgtagaataactattatattagacattgaatgaagcttaggaggtta 486  
Qy 499 tgaatgcgtatcactccttggacgtttactgacacacaccccttgcattcattcatt 558  
Db 487 tcaatgcctcaacaaactttatttccaaattccgataattatcattgacgtaaaaaatt 546  
Qy 559 tacacacaaaagatttgcctcccttgaaataatgcataataattatgctcatcattggaac 618  
Db 547 taacacacattcttttaatacagacagatcaggtgctgtggtgaatgtttcttctggtac 606  
Qy 619 aactggattcgtcgaaggttagtactagcagcattagacatagagcttaactataaagattcga 678  
Db 607 aactgggttctcgaagggagtcagtactaactcaacgaataattgtgcacgattcttcca 666  
Qy 679 tagcagggaatccatttatggcactcgtacgtggttccacaaacataattcttctccttagt 738

Db 667 tgaagaagactcactatttggtaaacgcaatcaatcccaacgacgacaattttaaaggtaat 726  
Qy 739 accgttccatcatcgtccttggaaatgtttactacattatcttactttgttagtagacttaa 798  
Db 727 accttccaccatggttttggatgatgaccacattagatacttacttcttggattccg 786  
Qy 799 ggttgtaattgtgaagaatttggggcgacacttttcttaaaaaaccatacagaattacaa 858  
Db 787 agttgttctaattgcacacgtttggaagaaaaactatttcaatcattacagaattacaa 846  
Qy 859 aatccocactattttagtggccctccctccttattgtgttttggctaaagaccattagt 918  
Db 847 agtggaaagtacttatttacttgaacacattaatgagcatttttgcataaagtgcaattag 906  
Qy 919 cgatacatcattatcagctcagcttaacaggaagttgctactggaagagctccttttaggaaa 978  
Db 907 tgaagaatcagatttattcgcctttaaagaaaattgcatctggtggtgcacatttatacaa 966  
Qy 979 agatgtcgacaagcagtagcaagaaggttgaataattacctggaatcatacagaagtatgg 1038  
Db 967 agaaatggggagatggtgaagaaaacggcttaaaacttgcaggcgaaggtatcgg 1026  
Qy 1039 attaaactgaacttgcgtcgttaattgatctaccctcctcataatgctgtgaaaaacaggttc 1098  
Db 1027 attaacagaacacacttcgctgttttatttaaccacgaacaaatgacgtcagacccggatc 1086  
Qy 1099 aactggaagaccccttgcatacatctaaagctaaagtttttagataacgctactggaagacc 1158  
Db 1087 aactggtaaaatagtagtaccatttcacgtgttaaagttgctacatcacaacaggaataat 1146  
Qy 1159 gctaggacccggagaaagagcgaataatgcttcaaaagtaaaatgattatgaagaagata 1218  
Db 1147 tttggggccaaatgaactggagaaatgtattttaaaggcgaatgataaaagaggtta 1206  
Qy 1219 ttacaacaatccggaagcaactattgatctactatgcacaaagatggttggcttcattctgg 1278  
Db 1207 ttataataatgaagaagctactaaagcaattatttaacaaagacggatggttgcgctctgg 1266  
Qy 1279 agatattggatattacgacgaagatggaatttcttatttagtgcgtactgagaagaact 1338  
Db 1267 tgattgtcttattatgacaatgatggccattttatttatttgacaggtggaagtcatt 1326  
Qy 1339 tattaaatcacagggatatacaggttgcctcgtgaactggaaaatcgtcttttacaaca 1398  
Db 1327 aattaaatataaaggttatcaggttgcacctcgtgaaaattgagggaaatactcttacaaca 1386  
Qy 1399 tccaagtattcgtcgtcgttacttacttggagttcccgacgaatttggtagacaattacc 1458  
Db 1387 tccgtatattgtgatgcggcgttacttggatataccggatgaagcgcggcgagcttcc 1446  
Qy 1459 tgcgtcgttgttggtagaattcggcaacgctgactgactgaaaaaggaagttcaagaattt 1518  
Db 1447 agctgcaggttctttagtacagactggaaaatatactaaacgacaaaatcgtacaaaatt 1506  
Qy 1519 tattgcagcacaagtcactccaacaaagcatcttcggagcgggtgctgatttggtagacg 1578  
Db 1507 tgtttccagtcgaagtctcaacagccaaatggctacgtggttgggttgaatttttggagaga 1566  
Qy 1579 tattccgaagggccctacttggaaaaactcatcagaaaaggagctccgagagaattttg 1634  
Db 1567 aattcccaaggaatcaactggaaaaatattgacagaaaagtgttaagacaaaattgttg 1622

RESULT 12  
AAAX32846  
ID AAAX32846 standard; DNA; 1639 BP.  
XX  
AC AAAX32846;  
XX  
DT 28-JUN-1999 (first entry)  
XX  
DE Mutant luciferase encoding DNA.











FT mutation transl\_except= (pos:887...889,aa:Phe)  
FT replace (150...153,AAAA)  
FT /\*tag= b  
FT mutation replace (278,T)  
FT /\*tag= d  
FT mutation replace (388,T)  
FT /\*tag= e  
FT mutation replace (554,A)  
FT /\*tag= f  
FT mutation replace (665..666,CA)  
FT /\*tag= g  
FT mutation replace (668..670,TGC)  
FT /\*tag= h  
FT mutation replace (792,C)  
FT /\*tag= i  
FT mutation replace (1064,G)  
FT /\*tag= j  
FT mutation replace (1068,C)  
FT /\*tag= k  
FT mutation replace (1202,A)  
FT /\*tag= l  
FT mutation replace (1632..1635,AATC)  
FT /\*tag= m  
FT mutation replace (1637..1639,AAGCTG)  
FT /\*tag= n

XX WO200120002-A1.

PN 22-MAR-2001.

XX 22-DEC-1999; 99WO-US30925.

XX 15-SEP-1999; 99US-0396154.

PR (PROM-) PROMEGA CORP.

XX Wood KV, Hall MP, Gruber M;

XX WPI; 2001-244809/25.

DR P-PSDB; AAU01222.

XX Mutant luciferases from Photuris pennsylvanica and Pyrophorus  
PT plagiophthalmus having increased thermostability, increased resistance  
PT to inhibitors and/or enhanced enzymological properties, useful for  
PT detecting ATP -

XX Disclosure: Fig 25; 260pp; English.

XX The sequence encodes a Beetle mutant luciferase, Luc49-7A5. Mutant  
CC luciferases from Photuris pennsylvanica and Pyrophorus plagiophthalmus,  
CC are created by recursive mutagenesis to have one or more desired  
CC properties, e.g. increased thermostability, increased resistance to  
CC inhibitors and/or enhanced enzymological properties relative to a  
CC wild-type enzyme. The mutant luciferases are useful for detecting ATP,  
CC for labeling a molecule, as genetic reporters, for immobilisation onto a  
CC solid surface, to produce hybrid proteins, for high temperature  
CC reactions, or for creating luminescent solutions. The luciferase enzymes  
CC have increased thermostability which allows storage of the enzyme without  
CC altering its activity and improves reproducibility and accuracy of  
CC assays.

XX Sequence 1639 BP; 555 A; 257 C; 323 G; 504 T; 0 other;

Query Match 31.9%; Score 563.6; DB 22; Length 1639;  
Best Local Similarity 59.8%; Pred. No. 3.4e-100;  
Matches 966; Conservative 0; Mismatches 644; Indels 6; Gaps 1;

Qy 25 aatgaagaagaacattagcgtggagcgtcctcgtgatagtcctcctggcgc 84

Db 7 aatgaagataaaattttatagcctgaacacattttatcctcctggcgtgatggac 66

Qy 85 ggcaggacaacattatcccaatcattgtataaaattgcattctttctctg-----aagc 138

Db 67 ggcaggacaacagatgttttcgcgcatattcctcgttatcgagatatttcaggatgcatagc 126  
Qy 139 aataatcgatgctcatacaaatgaagtaatacatcatctcaaatatttgaaacagcgtg 198  
Db 127 attgacaaatgctcatacaaagccctgttttatgaagagtttttaaaatttcgtg 186  
Qy 199 ccgcttagctgttagtagaacaatatggcttggaatgaacaaatgttgggtgtatg 258  
Db 187 tcgttttagcggaaagttttaaaagtatggaataaaacaaacgacacaatagcgtgtg 246  
Qy 259 cagtgaacaaataaaactttttaaactcctgctcctgctgtcttatacttaggaatacc 318  
Db 247 tagcgaatgggttgcaatttttccctataaattgcattcattctcttggataat 306  
Qy 319 agtagcaacatcaaatgatgtacacagatggagagtaactggtcattcttggaatatac 378  
Db 307 tgcagcacctgttagtgataaaatacatggaactgaattaatatacacagctcttgatgt 366  
Qy 379 aaacccaactatcgttttagttcaagaagacactcccgcttattcttgagagtcacagca 438  
Db 367 aaacccagcataatttttgcctcaagaatacttttcaaaaagctactgaaatgtaaaatc 426  
Qy 439 aaatctaagtttcattaaaaagtcgttagttatcgtatagcatgacatgacatttaagcgt 498  
Db 427 taaattaaaatgtagaacactattattatactgaacttaataagaagccttaggggtta 486  
Qy 499 tgaatgcgtatctactcctttgttcacgtttactgaccacacacctttgtccattgtcatt 558  
Db 487 tcaatgctcaacaactttatttccaaaattccgatatataatcttgacgttaaaaaatt 546  
Qy 559 tacaccaaagattttgatacccttgaaaaatcgacatttaattatgcatcatctcggaaac 618  
Db 547 taaacatatcttttaactgagacgacaggttcggttggttaagtgtttctctcgtgtac 606  
Qy 619 aactggattgcctaaaggttagtactgagccatagaagtcataactataaagattcgttca 678  
Db 607 aactgggttttcgaaggagtcactgaactcacagaataattgttcacagattttctat 666  
Qy 679 tagcaggatgccatttatggcaactcgtacggttcacaaacacataattcttctccttagt 738  
Db 667 tgaacaaagatcctactcttggtaacgcaattaatccacagacagcaattttaacggtta 726  
Qy 739 accgttccatcatgaccttggaaattttactacattatcttacttcttgtagtaggactaa 798  
Db 727 acccttccacattggttggtagatgacacacattagtagacttcttaacttggatctcg 786  
Qy 799 ggtgttaattggaataatttgaggcgcacattttcttaaaaccatacagaattacaa 858  
Db 787 agtgttctaattgcacacgcttggaagaaaaactattctcacatactatacagaattata 846  
Qy 859 aatccccactattgtagtgccctccagttatgggttttttggttaaaagcccttagt 918  
Db 847 agtggaagtacttactgtaccacaaactaatgaatggcatttttgcaaaagtcattagt 906  
Qy 919 cgatcaatcacgatttatcgagcttaacggaagtgtcactgagagagctccttttaggaaa 978  
Db 907 tgaagaagtcagatttatcgacatttaaaagaaattgcattcgttggcgacacctttatcaa 966  
Qy 979 agatgtccagaagcagtagcaaaagaggttgaaattactgggaatcatacacaaggtatgg 1038  
Db 967 agaaattggggagatggttgaaaaaacggttttaataacttgcaggcgaaggtatgg 1026  
Qy 1039 attaaactgaacctgtcgtgctgaatgtattaccctcatatgtcgtgtgaacacaggttc 1098  
Db 1027 attaacagaacacctcgtcgtgttttaattacacccgacaacatgacgtcagaccgggatc 1086  
Qy 1099 aactggaagacaccttgcacatacattaaagctaaaagttttagataacgctactcgggaagc 1158  
Db 1087 aactgtaaaatagtagcattttccacgctgttaaaagtgtcgtatcctacacaggaataat 1146  
Qy 1159 gctaggaccagaggaagagcgcaaatatgctttcaaatgtaaatgatttatgaaggata 1218

Db 1147 ttctggggccaaatgaactggagaattgtatttttaaaggcgacatgataatgaagggtta 1206  
QY 1219 ttacaacaatccggaggaactattgtactactattgacaaaagatggttggtcttcattctgg 1278  
Db 1207 ttataataatgaagaagctactaaagcaattatttaacaaagacggatggttgcgcctctgg 1266  
QY 1279 agatattggatattacgacgaagatggaaaatttctttatagttgatcgaattgaaaagaact 1338  
Db 1267 tgatatgtttattatgacaatgatggccattttatatatttggacaggtggaagtcatt 1326  
QY 1339 tattaaatacaaggatataatcaggttgcgcctgctgaactggaaaaatctgctttttacaaca 1398  
Db 1327 aattaaataaaagttatcaggttgcacctgctgaattggagggaatactcttacaaca 1386  
QY 1399 tccaagtattgctgcgggtgttactggaggtccggacgaatttggtagacaattacc 1458  
Db 1387 tccgtatatgttgatgcggcgttacttgggtataccggatgaagccggtggcgagcttcc 1446  
QY 1459 tgctgcttgtgtgtgttagaatctggcaagacgctgactgaaagggaagttcaagattt 1518  
Db 1447 agctgcagggtgtgtgtacagactggaaaaatctaaacgaacaaaatcgtaaaaaattt 1506  
QY 1519 tattgcagcacaaagtcactccaacaaaagcatcttcgagcggtgtcgtattttgtagacag 1578  
Db 1507 tgtttccagtcgaagtttcaacagccaaatggctacgtggtgggtgaaattttttggatga 1566  
QY 1579 tattccgaaagccctactgaaaaactcatcagaagaagctccgagaaaattttg 1634  
Db 1567 aattcccaaaaggatcaactggaaaaatttgacagaaaaagctttaagacaaaatgtttg 1622

Search completed: September 7, 2002, 22:32:12  
Job time: 9381 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 4, 2002, 15:04:01 ; Search time 13.45 seconds  
(without alignments)  
1568.933 Million cell updates/sec

Title: US-09-993-874-2  
Perfect score: 2812  
Sequence: 1 MESENHGRPRDIHVPGS.....KLIRKELREIFAORAPKSL 545

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 36719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1584.5	56.3	550	1 LUCI_PHOPY	P08659 photinus py
2	1491.5	53.0	548	1 LUCI_LUCCR	P13129 luciola cru
3	1458.5	51.9	548	1 LUCI_LUCLA	O01158 luciola lat
4	862	30.7	547	1 4CL1_TOBAC	O24145 nicotiana t
5	855	30.4	542	1 4CL1_TOBAC	O24146 nicotiana t
6	847	30.1	545	1 4CL1_SOLTU	P31684 solanum tub
7	838	29.8	545	1 4CL1_SOLTU	P31685 solanum tub
8	836.5	29.7	553	1 4CL1_VANPL	O24540 vanilla pla
9	833.5	29.6	563	1 4CL1_ORYSA	P17814 oryza sativ
10	824	29.3	562	1 4CL2_SOYBN	P31687 glycine max
11	820.5	29.2	556	1 4CL2_ARATH	O98725 arabidopsis
12	810	28.8	544	1 4CL2_PETCR	P14913 petroselinu
13	807.5	28.7	544	1 4CL1_PETCR	P14912 petroselinu
14	807	28.7	561	1 4CL1_ARATH	O42524 arabidopsis
15	773	27.5	537	1 4CL1_PINTA	P41636 pinus taeda
16	763.5	27.2	561	1 4CL3_ARATH	O98777 arabidopsis
17	719.5	25.6	569	1 4CL2_ORYSA	O42982 oryza sativ
18	638.5	22.7	293	1 4CL1_SOYBN	P31686 glycine max
19	615.5	21.9	560	1 LCFA_BACSU	P94547 bacillus su
20	535.5	19.0	561	1 LCFA_ECOLI	P29212 escherichia
21	444.5	15.8	562	1 LCFA_HAEIN	P46450 haemophilus
22	432	15.4	486	1 MENE_BACSU	P23971 bacillus su
23	420	14.9	469	1 MENE_LISMO	P58730 listeria mo
24	414	14.7	543	1 FAT2_YEAST	P38137 saccharomyc
25	406.5	14.5	548	1 YDID_ECOLI	P38135 escherichia
26	406	14.4	469	1 MENE_LISIN	O92ay8 listeria in
27	404	14.4	522	1 CAIC_ECOLI	P31552 escherichia
28	380.5	13.5	572	1 ACSA_BACSU	P39062 bacillus su
29	378	13.4	546	1 ALRK_PSEOL	Q00594 pseudomonas
30	372	13.2	699	1 LCFB_RAT	P18163 rattus norv
31	369.5	13.1	536	1 ENPE_ECOLI	P10378 escherichia
32	365	13.0	492	1 MENE_ETAAU	Q53634 staphylococ
33	350	12.4	699	1 LCFB_MOUSE	P41216 mus musculus

## ALIGNMENTS

### RESULT 1

ID	LUCI_PHOPY	STANDARD;	PRT;	550 AA.
AC	P08659;			
DT	01-JAN-1988 (Rel. 06, Created)			
DT	01-JAN-1988 (Rel. 06, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Luciferin 4-monooxygenase (EC 1.13.12.7) (Luciferase).			
OS	Photinus pyralis (North American firefly).			
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;			
OC	Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;			
OC	Elateriformia; Cantharoidea; Lampyridae; Photinus.			
OX	NCBI_TaxID=7054;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=87144243; PubMed=3821727;			
RA	de Wet J.R., Wood K.V., Deluca M., Helinski D.R., Subramani S.;			
RT	"Firefly luciferase gene: structure and expression in mammalian cells.";			
RL	Mol. Cell. Biol. 7:725-737(1987).			
RN	[2]			
RP	SUBCELLULAR LOCATION.			
RX	MEDLINE=87204117; PubMed=3554235;			
RA	Keller G.-A., Gould S., de Luca M., Subramani S.;			
RT	"Firefly luciferase is targeted to peroxisomes in mammalian cells.";			
RN	Proc. Natl. Acad. Sci. U.S.A. 84:3264-3268(1987).			
RN	[3]			
RP	X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).			
RX	MEDLINE=96398615; PubMed=8805533;			
RA	Conti E., Franks N.P., Brick P.;			
RT	"Crystal structure of firefly luciferase throws light on a superfamily of adenylate-forming enzymes.";			
RL	Structure 4:287-298(1996).			
RN	[4]			
RP	X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).			
RX	MEDLINE=99007339; PubMed=9788915;			
RA	Franks N.P., Jenkins A., Conti E., Lieb W.R., Brick P.;			
RT	"Structural basis for the inhibition of firefly luciferase by a general anesthetic.";			
RL	Biophys. J. 75:2205-2211(1998).			
CC	- FUNCTION: PRODUCES GREEN LIGHT WITH A WAVELENGTH OF 562 NM.			
CC	- CATALYTIC ACTIVITY: Photinus luciferin + O(2) + ATP - oxidized			
CC	Photinus luciferin + CO(2) + H(2)O + AMP + diphosphate + light.			
CC	- COPACITOR: MAGNESIUM.			
CC	- SUBCELLULAR LOCATION: Peroxisomal.			
CC	- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			

P09095 bacillus br  
O94049 candida alb  
P33121 homo sapien  
P41215 homo sapien  
O30409 b tyrocidin  
O68008 b bacitraci  
P40871 bacillus su  
O91d6 cavia porce  
O60011 kluyveromyc  
P14687 bacillus br  
P39518 saccharomyc  
O88813 rattus norv

34 349 12.4 1088 1 TYCA\_BACBR  
35 344 12.2 675 1 ACSA\_CANAL  
36 342 12.2 698 1 LCFB\_HUMAN  
37 335.5 11.9 699 1 LCFA\_HUMAN  
38 330.5 11.8 6486 1 TYCC\_BACBR  
39 329.5 11.7 6359 1 BACC\_BACLI  
40 328 11.7 539 1 DHBE\_BACSU  
41 323 11.5 698 1 LCFA\_CAVPO  
42 321.5 11.4 683 1 ACS1\_KLUOLA  
43 321 11.4 1098 1 GRSA\_BACBR  
44 320 11.4 744 1 LCPE\_YEAST  
45 319 11.3 683 1 LCPE\_RAT

```
DR EMBL; M15077; AAA29795.1; -
DR EMBL; X84848; CAA59283.1; -
DR EMBL; U03687; AAA03561.1; -
DR EMBL; U89934; AAB64396.1; -
DR EMBL; U89935; AAB64399.1; -
DR PIR; A26772; A26772.
DR PDB; ILIC1; 26-MAR-97.
DR PDB; 1BA3; 11-NOV-98.
DR InterPro; IPR000873; AMP-bind.
DR Pfam; PF00501; AMP-binding; 1.
DR PROSITE; PS00342; MICROBODIES_CTER; 1.
DR PROSITE; PS00455; AMP_BINDING; 1.
KW Oxidoreductase; Monooxygenase; Photoprotein; Luminescence; Magnesium;
FT Peroxisome; 3D-structure.
FT SITE 548 550 MICROBODY TARGETING SIGNAL.
SQ SEQUENCE 550 AA; 60745 MW; E380FCE9D56ACCDE CRC64;

Query Match 56.3%; Score 1584.5; DB 1; Length 550;
Best Local Similarity 55.5%; Pred. No. 4.5e-107;
Matches 305; Conservative 103; Mismatches 137; Indels 5; Gaps 3;

QY 1 MEE-ENIRHGERPRDIVHPGAGQOLYQSLYKFAFPE--AIIDAHTNEVISYAOIFETS 57
DB 1 MEDAKNIKKGPAPFYLEDGTAGEQLHKAMKRYALVPCTIAFTDAHIEVNTIYAEYFEMS 60
QY 58 CRLAVSIEQYGLNENNVGVCSENNINFPVLAALYLGIPVATSDMYTDCGLTGHINI 117
DB 61 VRLAEMKRYGLNTHIRVWCSENSLQFMPVLGALFTGVAVAPANDIYNRELNSMI 120
QY 118 SKPTIMFSKKALPILRVQNLSPFKVVDSDMYDINGVECVSTVARYTDHTFDPLS 177
DB 121 SQTVPVSVSKKGLQILNVQKLPITQIIIMDSKTDYOGFQSGMYTFTVSHLPPGFNEYD 180
QY 178 FPPKDFDPLEKALINSSSGTGLPKGVVLSHRSITIRFVHSRDIYCTRTVPQTSLSL 237
DB 181 FVPESFDRDKTALINSSSGTGLPKGVVLSHRSITIRFVHSRDIYCTRTVPQTSLSL 240
QY 238 VPFHAFGMFTTSLYFVVGKLVYMLKKEGALFLKTIONYKIPTTIWAPPVNVFLAKSPL 297
DB 241 VPFHGFGMFTTGLYLCGFRVVLMYRFEELFLSLQDYKIQSALLVPTLFSFAKSTL 300
QY 298 VDQYDLSSLETVATGACPLGKDVAEAVAKRLKPLGIQYGLTETCCAVMITPHNAVKTG 357
DB 301 IDKYDLSNLHEIATSGAPLSKEVGEAVAKRPHLPGLRGYGLTETTSAILITPEGDKPG 360
QY 358 STGRPLPYIAKAVLDNATGKALGPERGEICQFQSEIMKGYNNPEATIDTDKQGLHS 417
DB 361 AVGVVVPFPEAKVVDLDTGKTLGVNQRGELCYRGPIMSGVYNNPEATNALIDKQGLHS 420
QY 418 GDIGYDDEGNFFIVDRKLKELIKYGYOVAPAELENLLLOHPSIADAGVTGVPDEFQGL 477
DB 421 GDIAWDEDEHFFIVDRKLSLKYGYOVAPAELESILLQHPNIFDAGVAGLDDADGL 480
QY 478 PAACVWLESGKTLTEKEVQDFTAAQVPTPKHLRGVGVFVDSIPKGTKLIRKELREIF- 536
DB 481 PAAVVWLEHGKMTKEIYDVASQVTTAKKLRGGVGVFVDEVPKGLTGKLDARKITREIL 540
QY 537 -AQRAPKSL 545
DB 541 KAKKGKSKL 550

RESULT 2
ID LUCCR
AC LUCCR
DT 01-JAN-1990 (Rel. 13; Created)
DT 01-JAN-1990 (Rel. 13; Last sequence update)
DT 16-OCT-2001 (Rel. 40; Last annotation update)
DE Luciferin 4-monooxygenase (EC 1.13.12.7) (Luciferase).
OS Luciola cruciata (Japanese firefly) (Genji firefly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
```

```
OC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
OC Elateriformia; Cantharoidea; Lampyridae; Luciola.
RN NCBI_TaxID=7051;
RP SEQUENCE FROM N.A. PubMed=2473944;
RX MEDLINE=89326143; PubMed=2473944;
RT Masuda T., Tatsumi H., Nakano E.;
RT "Cloning and sequence analysis of cDNA for luciferase of a Japanese
RT firefly, Luciola cruciata.";
RL Gene 77:265-270(1989).
CC -!- FUNCTION: PRODUCES GREEN LIGHT WITH A WAVELENGTH OF 544 NM.
CC -!- CATALYTIC ACTIVITY: Photinus luciferin + O(2) + ATP -> oxidized
CC Photinus luciferin + CO(2) + H(2)O + AMP + diphosphate + light.
CC -!- COFACTOR: MAGNESIUM.
CC -!- SUBCELLULAR LOCATION: Peroxisomal (By similarity).
CC -!- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M26194; AAA29135.1; -
DR PIR; JS0181; JS0181.
DR HSP; P08659; IBA3.
DR InterPro; IPR000873; AMP-bind.
DR Pfam; PF00501; AMP-binding; 1.
DR PROSITE; PS00342; MICROBODIES_CTER; 1.
DR PROSITE; PS00455; AMP_BINDING; 1.
KW Oxidoreductase; Monooxygenase; Photoprotein; Luminescence; Magnesium;
KW Peroxisome.
FT SITE 546 548 MICROBODY TARGETING SIGNAL (POTENTIAL).
SQ SEQUENCE 548 AA; 60017 MW; 2052D6189E79109F CRC64;

Query Match 53.0%; Score 1491.5; DB 1; Length 548;
Best Local Similarity 53.1%; Pred. No. 2.4e-100;
Matches 289; Conservative 98; Mismatches 154; Indels 3; Gaps 2;

QY 3 ENIRHGERPRDIVHPGAGQOLYQSLYKFAFPE-AIIDAHTNEVISYAOIFETSCRLA 61
DB 7 DENIVVGPKFPPIEGSGAGTQLRYMERYAKLGAIAETNAVTVGYDYSYAEVLEKSCCLG 66
QY 62 VSIEQYGLNENNVGVCSENNINFPVLAALYLGIPVATSDMYTDCGLTGHLSKPT 121
DB 67 KALQNYGLVVDGRIALCSENCSEFFIPVIAGLFVGAVPTNEIYTLRELHSLGSKPT 126
QY 122 IMFSSKKALPILRVQNLSPFKVVDSDMYDINGVECVSTFVARYTDHTFDPLSFTPK 181
DB 127 IVFSSKKGLDKVITQKTTIKTIVILDSKVDRYGCQCLDTFIKRNTPPGQASSFKTV 186
QY 182 FDPDLEKALINSSSGTGLPKGVVLSHRSITIRFVHSRDIYCTRTVPQTSILSLVPH 241
DB 187 EYDRKEQVALINSSSGTGLPKGVLTHTNTVTRSHARDPIYGNQSPGTAVLIVVPH 246
QY 242 HAFGMFTTSLYFVVGKLVYMLKKEGALFLKTIONYKIPTIIVVAPPVNVFLAKSPLVDQY 301
DB 247 HGFGMFTTGLYLCGFRVVMKTKDEETFLKTLQDYKCTSVILVPTLFAILNKSELLNKY 306
QY 302 DLSSLETVATGACPLGKDVAEAVAKRLKPLGIQYGLTETCCAVMITPHNAVKTGSGTGR 361
DB 307 DLSNLVEIATSGGAPLSKEVGEAVARRFNLPGVRQYGLTETTSAILITPEGDDKFGASK 366
QY 362 PLPYIAKAVLDNATGKALGPERGEICQFQSEIMKGYNNPEATIDTDKQGLHSGDGI 421
DB 367 VVPLFAKAVIDLDTKSLGPNRRGEVCVKGPLMKGYNNPEATKELIDEEGLHTGDI 426
QY 422 YYDEGNFFIVDRKLKELIKYGYOVAPAELENLLLOHPSIADAGVTGVPDEFQGLPAAC 481
DB 422 YYDEGNFFIVDRKLKELIKYGYOVAPAELENLLLOHPSIADAGVTGVPDEFQGLPAAC 481
```

```
Db 427 YYDEKHEFFIVDRLSLKIKYQVPPAELESVLLQHSIFDAGVAGVDPDPVAGELPGAV 486
Qy 482 VVLESGSKTLTEKEVODFAAQVPTPKHLRGVGVFVDSIPKPGTGKILIRKELREIFAQAP 541
Db 487 VVLESGKNTKEVMDYVVASQVSNKRLRGVVRFEVDEVPKGLTGKIDGRAIREIL--KKP 544
Qy 542 KSKL 545
Db 545 VAKM 548

RESULT 3
LUCI_LUCIA STANDARD; PRT; 548 AA.
ID LUCI_LUCIA
AC Q01158;
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Luciferin 4-monooxygenase (EC 1.13.12.7) (Luciferase).
OS Luciola lateralis (Firefly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
OC Elateriformia; Cantharoidae; Lampyridae; Luciola.
OX NCBI_TaxID=7052;
RN [1]
RP SEQUENCE FROM N.A. PubMed-1610896;
RX MEDLINE-92305054; PubMed-1610896;
RA Tatsumi H., Kajiyama N., Nakano E.;
RT "Molecular cloning and expression in Escherichia coli of a cDNA clone
RL encoding luciferase of a firefly, Luciola lateralis.";
RL Biochim. Biophys. Acta 1131:161-165(1992).
CC -!- FUNCTION: PRODUCES GREEN LIGHT.
CC -!- CATALYTIC ACTIVITY: Photinus luciferin + O(2) + ATP -> oxidized
CC Photinus luciferin + CO(2) + H(2)O + AMP + diphosphate + light.
CC -!- COPACITOR: MAGNESIUM.
CC -!- SUBCELLULAR LOCATION: Peroxisomal (By similarity).
CC -!- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME
CC FAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; X66919; CAA47358.1; -.
DR PIR; S23437; S23437.
DR HSP; P08659; IBA3.
DR InterPro; IPR000873; AMP-bind.
DR Pfam; PF00501; AMP-binding; 1.
DR PROSITE; PS00342; MICROBODIES_CTER; 1.
DR PROSITE; PS00455; AMP_BINDING; 1.
KW Oxidoreductase; Monooxygenase; Photoprotein; Luminescence; Magnesium;
KW Peroxisome.
FT SITE 546 MICROBODY TARGETING SIGNAL (POTENTIAL).
SQ SEQUENCE 548 AA; 60125 MW; AC62F9320BB6D4A6 CRC64;

Query Match 51.9%; Score 1458.5; DB 1; Length 548;
Best Local Similarity 51.8%; Pred. NO. 5.8e-98;
Matches 282; Conservative 101; Mismatches 158; Indels 3; Gaps 2;

Qy 3 ENIRHGERPRDIVHPSGAGQOLYSLKFASFPE-AIIDAHTNEVISYAOIFETSCRLA 61
Db 7 DENIVGPEFFPIEGSGAGALRYMDRYAKLGAIAFTNALTGYDVTYAEYLEKSCCLG 66
Qy 62 VSIEQYGLNNVGVCSENNINFFNPVLAALYLGIPVATNSNDMTDGLTGLHLNISKPT 121
Db 67 EALKNYGLVWDCRIALCSENCCEFFIPVLGLFVGIVGVAPTNEITLRELHSLGISKPT 126
Qy 122 IMFSSKKALPLILRVQOONLSFIKKVWVDSMYDINGVECVSTFVARYTDHTFDPLSFTPK 181
```

```
Db 127 IVFSSKKGLDKVITVQKTTVTAIKTIVILDSKVDYRGYQSMDFIRKNTPQGFKGSFKTV 186
Qy 182 DFDPLEKIALINSSGTTGLPKGVVLSHSLTIRVHSDRDPITGYTRTVPTQTSILSLVPFH 241
Db 187 EVNRKEQVALINSSSGTGLPKGVQLTHENAVTRSHARDPIYGNVSGTALLVVPFH 246
Qy 242 HAFGMFTTSLYFVVGKLVVMLKKFEGALFLKTIQNKPIITVIVAPPVWVFLAKSLVDQY 301
Db 247 HGFGMFTTGLYTCGFRIVMLTKFDEETFLQDYKSSVILVPTFLFALNRSELLDKY 306
Qy 302 DLSLSTEATGAPLGGKDVAAVAKRLKPLGIIOGVLGTETCCAVMITPHNAVKTGSTGR 361
Db 307 DLSNLVEIASGGAPLSKEIGEAARFNLPGVRQGTGLTETTSATIIITPEGDDKPCASGR 366
Qy 362 PLPIYIKAVLDNATGKALGPGERGEICFQSEIMKGYNNPEATIDTIDKQWMLHSGDIG 421
Db 367 VVPLPKAKVIDLTKTLGPNRRGVCVKGPMLMGVYDNPENATREIIDEQWMLHTGDIG 426
Qy 422 YYDEGNCFFIVDRLEKELIKYQVAPAELENLLQHSIADAGVTGVPDEFGGQLPAAC 481
Db 427 YYDEKHEFFIVDRLSLKIKYQVPPAELESVLLQHPNIFDAGVAGVDPDPVAGELPGAV 486
Qy 482 VVLESGSKTLTEKEVODFAAQVPTPKHLRGVGVFVDSIPKPGTGKILIRKELREIFAQAP 541
Db 487 VVLEKGSMTKEVMDYVVASQVSNKRLRGVVRFEVDEVPKGLTGKIDGRAIREIL--KKP 544
Qy 542 KSKL 545
Db 545 VAKM 548

RESULT 4
4CL1_TOBAC STANDARD; PRT; 547 AA.
ID 4CL1_TOBAC
AC 024145;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 4-coumarate--CoA ligase 1 (EC 6.2.1.12) (4CL 1) (4-coumaroyl-CoA
DE synthase 1).
GN 4CL1
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-96416441; PubMed-8819324;
RA Lee D., Douglas C.J.;
RT "Two divergent members of a tobacco 4-coumarate:coenzyme A ligase
RT (4CL) gene family, cDNA structure, gene inheritance and expression,
RT and properties of recombinant proteins.";
RL Plant Physiol. 112:193-205(1996).
CC -!- CATALYTIC ACTIVITY: ATP + 4-coumarate + CoA -> AMP + diphosphate +
CC 4-coumaroyl-CoA.
CC -!- PATHWAY: BRANCH-POINT REACTIONS BETWEEN GENERAL PHENYLPROPANOID
CC METABOLISM AND PATHWAYS LEADING TO VARIOUS SPECIFIC END PRODUCTS.
CC -!- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME
CC FAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; U50845; AAB18637.1; -.
DR HSP; P08659; ILCI.
DR Mendel; 24477; Nicta;1179;24477.
```





```
Db 306 RTVMGAAPLGRKELEDAV--RAKFNNAKLGQGYGTEAGPVLAMCLAFAKEPFD-IKSGA 362
QY 359 TGRPLPYIAKAVLDNATKALGPGRGEICFQSEIMKGYNNPEATIDTIDKQGLHSG 418
Db 363 CQTVVNAEMKIVDPDTCGLPRNQPGEICIRGDOIMKGLNDPEATARTIEKGNLHTG 422
QY 419 DIGYDEGDNFFIVDLRLKELIKYQVAPAELEMLLQHPHSIADAGVGVDFEGGOLP 478
Db 423 DIGFIDDDDELFIIDLRLKELIKYQVAPAELEALLNHDPDISDAVVPMIDEQAGEVP 482
QY 479 ACNVLESCKTITEKQVDFIAAQTPTPKHLRGVGVFVDSIPKGTGKILREL 533
Db 483 VAFVVRKSNSTIDEVKDFISKQVIFYKRIK-RVFFVETVPKSPGKILRKDLR 536

RESULT 8
4CLL_VANPL STANDARD: PRT: 553 AA.
AC 024540;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 4-coumarate--CoA ligase (EC 6.2.1.12) (4CL) (4-coumaroyl-CoA
synthase).
GN 4CL.
OS Vanilla planifolia (Vanilla).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Orchidaceae;
OC Vanilloid clade; Vanillinae; Vanilla.
OX NCBI_TaxID=51239;
RN [1]
RP SEQUENCE FROM N.A.
RA Brodelius P., Xue Z.T.;
RT "Isolation and characterization of a cDNA from cell suspension
RT cultures of Vanilla planifolia encoding 4-coumarate: coenzyme A
RT ligase.";
RL Plant Physiol. Biochem. 35:497-506(1997).
CC -|- CATALYTIC ACTIVITY: ATP + 4-coumarate + CoA = AMP + diphosphate +
CC 4-coumaroyl-CoA.
CC -|- PATHWAY: BRANCH-POINT REACTIONS BETWEEN GENERAL PHENYLPROPANOID
CC METABOLISM AND PATHWAYS LEADING TO VARIOUS SPECIFIC END PRODUCTS.
CC -|- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME
CC FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL: X75542; CAA53230.1; -.
CC HSSP: P08659; ILCI.
CC InterPro: IPR000873; AMP-bind.
CC Pfam: PF00501; AMP-binding; 1.
CC PROSITE: PS00455; AMP-BINDING; 1.
KW Ligase; Phenylpropanoid metabolism.
SQ SEQUENCE 553 AA; 60095 MW; 9A2D250BC84EA1CB CRC64;

Query Match 29.7%; Score 836.5; DB 1; Length 553;
Best Local Similarity 35.9%; Pred. No. 5.4e-53;
Matches 198; Conservative 102; Mismatches 177; Indels 75; Gaps 14;

QY 26 YOSLYKFAFPAAIDAHTNEVISAQIFETSCRVAISYEQVGLNENNVGVCSNNINF 85
Db 36 FENISKFSRP-CLINGATDEFTADVDELISRRVSGSLKLGKIQGDTIMILLNSPEF 94
QY 86 FNPVLAALYLGIPVATSDNMYTDGELTGHNLNISKPTIFESSKKALPLILRVQONLSFIKK 145
Db 95 VFAFLGASFGISITMANPFTSTEVIKQAKASNAKLI-----ITQ 135
```

```
QY 146 VVVIDSMYDI---NGVECVSTFVARYTDTHTFDPLSF---TPKDFDPLEKIAL-----I 192
Db 136 GCVYDKVKDYACENGKVIISIDTTTTTADDAANILHFSELTSAGADENEMPKVEISPDGVVAL 195
QY 193 MSSSTGTTGLPGKGVLSHRSLTIR-----FVHSRDPYIGTRVTPQTSILSLVPF 240
Db 196 PYSSGTTGLPGKVMLTRKGLVTSVAQQVDGENPNLYMHSDD-----VLLCVLPL 244
QY 241 HHAFGMFTTSLVFVVGK---VVMKKFEGALFKLTQNYKIPTIVVAPPVVMVFLAKSP 296
Db 245 PHYSLSNVL---LCGLRAGSGILIMQKPEIVPFVLELQKVKVTIGTPVPIVLAIAKST 301
QY 297 LVQDYDLSLSEVATGCGAPLGRKDAEAVAKRLKLPG--IIQGYGLTEN-----CCAVMI 348
Db 302 VVDNYDLSVVRTVMSSGAAPLCKELEDAV--RAKFPNNAKLGQGYGTEAGPVLAMCLAFAP 359
QY 349 TPNNAVKTGTCGRPLPYIAKAVLDNATKALGPGRGEICFQSEIMKGYNNPEATIDT 408
Db 360 EPFD-IKSGACGTVVNAEMKIVDPETGSSLPNRNHPGEICIRGDOIMKGLNDPEATART 418
QY 409 IDKQGLHSGDIGYDEGDNFFIVDLRLKELIKYQVAPAELEMLLQHPHSIADAGVGTG 468
Db 419 IDKQGLHSGDIGYDEGDNFFIVDLRLKELIKYQVAPAELEMLLQHPHSIADAGVGTG 478
QY 469 VDFEGGOLPAACVVLVLESKTLITEKQVDFIAAQTPTPKHLRGVGVFVDSIPKGTGKLI 528
Db 479 MKDEAAGEVPVAFVVKSNHGNHTEDEIKQFISKQVIFYKRI-NRVFFVEATPKAPSGKIL 537
QY 529 RKELREIFAQRA 540
Db 538 RKDLRLAALAAA 549

RESULT 9
4CLL_ORYSA STANDARD: PRT: 563 AA.
AC P17814;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE 4-coumarate--CoA ligase 1 (EC 6.2.1.12) (4CL 1) (4-coumaroyl-CoA
synthase 1).
GN 4CL1 OR 4CL.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV JAPONICA;
RX MEDLINE=91045096; PubMed=2235510;
RA Zhao Y., Kung S.D., Dube S.K.;
RT "Nucleotide sequence of rice 4-coumarate:CoA ligase gene, 4-CL.1.";
RL Nucleic Acids Res. 18:6144-6144(1990).
CC -|- CATALYTIC ACTIVITY: ATP + 4-coumarate + CoA = AMP + diphosphate +
CC 4-coumaroyl-CoA.
CC -|- PATHWAY: BRANCH-POINT REACTIONS BETWEEN GENERAL PHENYLPROPANOID
CC METABOLISM AND PATHWAYS LEADING TO VARIOUS SPECIFIC END PRODUCTS.
CC -|- INDUCTION: BY FUNGAL ELICITOR AND UV IRRADIATION.
CC -|- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME
CC FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL: X52623; CAA36850.1; -.
CC
```



\*Molecular cloning and expression of 4-coumarate:coenzyme A ligase, an enzyme involved in the resistance response of soybean (Glycine max L.) against pathogen attack.\*  
 RL Plant Physiol. 102:1147-1156(1993).  
 CC -|- CATALYTIC ACTIVITY: ATP + 4-coumarate + CoA -> AMP + diphosphate + 4-coumaroyl-CoA.  
 CC -|- PATHWAY: BRANCH-POINT REACTIONS BETWEEN GENERAL PHENYLPROPANOID METABOLISM AND PATHWAYS LEADING TO VARIOUS SPECIFIC END PRODUCTS.  
 CC -|- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: X69955; CAC36095.1; -  
 CC PIR: S31709; S31709.  
 CC PIR: P00772; P00772.  
 CC HSSP: P08659; ILCT.  
 CC InterPro: IPR000873; AMP-bind.  
 CC Pfam: PF00501; AMP-binding; 1.  
 CC PROSITE: PS00455; AMP\_BINDING; 1.  
 CC KW Ligase; Phenylpropanoid metabolism; Multigene family.  
 CC SQ SEQUENCE 562 AA; 60963 MW; 2AB4652DDA5160B2 CRC64;  
 -----  
 Query Match 29.3%; Score 824; DB 1; Length 562;  
 Best Local Similarity 36.5%; Pred. No. 4.4e-52;  
 Matches 198; Conservative 100; Mismatches 192; Indels 52; Gaps 13;  
 -----  
 QY 26 YQLYKFPASPEAIIDAHTNEVISVAQIFETSCRLAVSIEQYGLNENNVGVCSENNIN 85  
 DB 47 FQNLQFAHRP-CLIVGPASKTFTYADTHLISKTAAGLSNGLIUKDQVWMLLQNSADF 105  
 QY 86 FNPVLAALYIGIPVATSDMYTGDGELTGHNLISKPTIMFSSKKALPLILRVQOONLSFIKK 145  
 DB 106 VFSLAISMIGAVATTANPFYTAPEIFKQFTVSKAKLI-----ITQ 146  
 QY 146 VVVDYDNGVINGECVSTF-VARYTDHTDPLSFT-----PK-DFDPLEKIALIM 193  
 DB 147 AMYVDKLRNHDGAKLGEDEKVVTVDDPPENCLEHFSVLSSEANESDVPEVEIHPDDAVAMPF 206  
 QY 194 SSGTGTGLPKGVVLSHRSLTIRFVHSRD---PIYGRTRVPOTSILSLVPPHFAFGMTT 249  
 DB 207 -SSGTTGLPKGVILTTHKSULTTVAQQVDGENPNLYLT---TEDVLLCVLPUPHFISLNSV 262  
 QY 250 LSYFV-VGLKVVMLKFEKALFKTIQNTKIPTIVVAPPMVFLAKSPVLDQYDLSLSL 308  
 DB 263 LLCAIRAGSAVILMOKFEIGTLLLIQRHRSVAMVVPVPLVLAALAKNPVADFDLSIRL 322  
 QY 309 VATGGAPLKDVAEAVAKRLKLPGLIIQGYGLTET-----CAVMITPHNAVKTGSTRP 362  
 DB 323 VLGAAPLPGLEELALRNMPQAVLQGGYGMTEAGPVLSCMLGFAKQFPQ-TKSGSCGT 381  
 QY 363 LPYIKAKVLDNATGKALGPGERGEICFQSEMINKGYNNPEATIDTDKDWLHSGDYG 422  
 DB 382 VRNAELKVVDPETGRSLGYNQGEICIRGQIMKGYLNDEAATASTIDSEGLHTGDVGY 441  
 QY 423 YDEGDNFFIVDRLEKILKYGVQVAPAELENLLQHPSTIADAGVGVDPDFEGQLPAACV 482  
 DB 442 VDDDEIFIVDRYKELIKYGVQVPAELGLLVSHPSIADAAVVPQDVAAGEVPVAFV 501  
 QY 483 VLSEGKTLTEKEVQDFTAAQVTPTKHLRGVGVFVSDIPKPGTGKILRKLRL---EIFQR 539  
 DB 502 VRNGFDLTEEAQVKEFTAKQVVFYKRLH-KVYFVHAIPKSPSGKILRDLKRAKLTAAQ 560  
 QY 540 AP 541  
 DB 561 TP 562



CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; X13325; CAA31697.1; -;  
DR EMBL; X05351; CAA28960.1; -;  
DR PIR; S01667; S01667;  
DR PIR; S15695; S15695;  
DR HSSP; P08659; ILCI.  
DR InterPro; IPR000873; AMP-bind.  
DR Pfam; PF00501; AMP-binding; 1.  
DR PROSITE; PS00455; AMP\_BINDING; 1.  
KW Ligase; Phenylpropanoid metabolism; Multigene family.  
SQ SEQUENCE 544 AA; 59783 MW; B477965C68F8C534 CRC64;

Query Match 28.8%; Score 810; DB 1; Length 544;  
Best Local Similarity 34.8%; Pred. No. 4.3e-51;  
Matches 189; Conservative 105; Mismatches 167; Indels 82; Gaps 13;  
QY 39 IIDAHTNEVISVAQIFETSCRLAVSIEQYGLNENNVGVCSENNINFFNPVLAALYLGTIP 98  
DB 45 LINGATGETFTYSQVVELLSRKVASGLNKLGIOGGDTIMLLPNSPEYFAFLGASRGAI 104  
QY 99 VATSNDMYTDELTHLNLISKPTIMFSSKKALPLILRVQONLSFIKKVVDMSYDINGV 158  
DB 105 STMANPFTSAEVIKQLKAS-----LAKLIITQACY---V 136  
QY 159 ECVSTFVAR-----YTDHTFDPLSFT-----PKDFPLEKIALINSSSGTTGL 201  
DB 137 DKVKDYAAEKNTIICIDDAPODCLHFKLMEADSEMPVVDSDVVALPYSSGTTGL 196  
QY 202 PKGVVLSHRSLTIR-----FVHSRDPYIGTRVTPOTSILSLVPFHFAFGMFTT 249  
DB 197 PKGVMTHKGLVTSVAQQVGDGNPNLYMHSD-----VMICILPLFIHYSLNAY 245  
QY 250 LSYFV-VGLKVVMLKKFEGALFKTIQNYKIPTIVVAPVPMVFLAKSPLDVQYDLSLSTE 308  
DB 246 LCCGLRAGVTIILMOKFDIVPPELEIQYKVTIGFPVPIVLAIAKSPVVDKYLSSVRT 305  
QY 309 VATGAPGLKGVAAVAKRLKLPG--IIQGYGLTET-----CCAVMTIPHNATVGTSGT 360  
DB 306 VMSGAPLGLKELEDVAV--RAKFPNAKLGQGYGWTAGPVLAMCLAFAPKEPYE-IRSGACG 362  
QY 361 RPLPIKAVKVLNATGKALGPCERGEICFQSEMIKMGYNPNPEATIDTIDKQWLHSGDI 420  
DB 363 TVVRAEMKVIDPETAFLPRNGRQCEIRGQIMKGLNDPESTRITIDEEGWLHTGDI 422  
QY 421 GYDDEGFFIVDLRELKIKYGVAPAELENLLQHPSTADAGVTGVPDFFGQLPAA 480  
DB 423 GPIDDDDELIVDLRLKEIKIKYGFQVAPAELEALLTHTPTISDAVVPWIDEKAGEVPA 482  
QY 481 CVVLESGTLEKEVODFTAAQVTPTKHLRGVGVFVDSIPKGPOTKGLIRKELR-EIFAQR 539  
DB 483 FVVRTNGFTTTEEIKQFVSKQWVFYKRI-FRVFFVDAIPKSPGKILRKDLRAKIASGD 541  
QY 540 APK 542  
DB 542 LPK 544

RESULT 13  
4CL1\_PETCR STANDARD; PRT; 544 AA.  
AC P14912;  
DT 01-APR-1990 (Rel. 14, Created)  
DT 01-APR-1990 (Rel. 14, Last sequence update)  
DT 01-MAR-2002 (Rel. 41, Last annotation update)  
DE 4-coumarate-CoA ligase 1 (EC 6.2.1.12) (4CL 1) (4-coumaroyl-CoA  
DE synthase 1).  
GN 4CL1 OR 4CL-1.  
OS Petroselinum crispum (Parsley) (Petroselinum hortense).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

CC Asteridae; euasterids II; Apiales; Apiaceae; Petroselinum.  
OX NCBI\_TaxID=4043;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=89005119; PubMed=3169018;  
RA Lozoya E., Hoffmann H., Douglas C., Schulz W., Scheel D.,  
RA Hahlbrock K.;  
RT "Primary structures and catalytic properties of isoenzymes encoded by  
RT the two 4-coumarate:CoA ligase genes in parsley.";  
RL Eur. J. Biochem. 176:661-667(1988).  
RN [2]  
RP SEQUENCE OF 1-8 FROM N.A.  
RA Douglas C., Hoffmann H., Schulz W., Hahlbrock K.;  
RT "Structure and elicitor or U.V.-light-stimulated expression of two  
RT 4-coumarate:CoA ligase genes in parsley.";  
RL EMBL J. 6:1189-1195(1987).  
CC -|- CATALYTIC ACTIVITY: ATP + 4-coumarate + CoA = AMP + diphosphate +  
CC 4-coumaroyl-CoA.  
CC -|- PATHWAY: BRANCH-POINT REACTIONS BETWEEN GENERAL PHENYLPROPANOID  
CC METABOLISM AND PATHWAYS LEADING TO VARIOUS SPECIFIC END PRODUCTS.  
CC -|- INDUCTION: BY FUNGAL ELICITOR AND UV IRRADIATION.  
CC -|- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME  
CC FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X13324; CAA31696.1; -;  
DR EMBL; X05350; CAA28959.1; -;  
DR PIR; S01667; S01667.  
DR HSSP; P08659; IBA3.  
DR InterPro; IPR000873; AMP-bind.  
DR Pfam; PF00501; AMP-binding; 1.  
DR PROSITE; PS00455; AMP\_BINDING; 1.  
KW Ligase; Phenylpropanoid metabolism; Multigene family.  
SQ SEQUENCE 544 AA; 59825 MW; 22BBAD78F25D0C8 CRC64;

Query Match 28.7%; Score 807.5; DB 1; Length 544;  
Best Local Similarity 34.9%; Pred. No. 6.6e-51;  
Matches 186; Conservative 103; Mismatches 163; Indels 81; Gaps 12;  
QY 39 IIDAHTNEVISVAQIFETSCRLAVSIEQYGLNENNVGVCSENNINFFNPVLAALYLGTIP 98  
DB 45 LINGATGETFTYSQVVELLSRKVASGLNKLGIOGGDTIMLLPNSPEYFAFLGASRGAI 104  
QY 99 VATSNDMYTDELTHLNLISKPTIMFSSKKALPLILRVQONLSFIKKVVDMSYDINGV 158  
DB 105 STMANPFTSAEVIKQLKASQ-----AKLIITQACY---V 136  
QY 159 ECVSTFVAR-----YTDHTFDPLSFT-----PKDFPLEKIALINSSSGTTGL 201  
DB 137 DKVKDYAAEKNTIICIDDAPODCLHFKLMEADSEMPVVDSDVVALPYSSGTTGL 196  
QY 202 PKGVVLSHRSLTIR-----FVHSRDPYIGTRVTPOTSILSLVPFHFAFGMFTT 249  
DB 197 PKGVMTHKGLVTSVAQQVGDGNPNLYMHSD-----VMICILPLFIHYSLNAY 245  
QY 250 LSYFV-VGLKVVMLKKFEGALFKTIQNYKIPTIVVAPVPMVFLAKSPLDVQYDLSLSTE 308  
DB 246 LCCGLRAGVTIILMOKFDIVPPELEIQYKVTIGFPVPIVLAIAKSPVVDKYLSSVRT 305  
QY 309 VATGAPGLKGVAAVAKRLKLPG--IIQGYGLTET-----CCAVMTIPHNATVGTSGT 360  
DB 306 VMSGAPLGLKELEDVAV--RAKFPNAKLGQGYGWTAGPVLAMCLAFAPKEPYE-IRSGACG 362  
QY 361 RPLPIKAVKVLNATGKALGPCERGEICFQSEMIKMGYNPNPEATIDTIDKQWLHSGDI 420



```
DE synthase).
GN 4CL.
OS Pinus taeda (loblolly pine).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferales; Pinaceae; Pinus.
OX NCBI_TaxID=3352;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=L.; TISSUE=Xylem;
RA Voo K.S.;
RL Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Needle;
RX MEDLINE=97161117; PubMed=9008388;
RA Zhang X.H., Chiang V.L.;
RT "Molecular cloning of 4-coumarate:coenzyme A ligase in loblolly pine
RT and the roles of this enzyme in the biosynthesis of lignin in
RT compression wood.";
RL Plant Physiol. 113:65-74(1997).
CC -|- CATALYTIC ACTIVITY: ATP + 4-coumarate + CoA -> AMP + diphosphate +
CC 4-coumaroyl-CoA.
CC -|- PATHWAY: BRANCH-POINT REACTIONS BETWEEN GENERAL PHENYLPROPANOID
CC METABOLISM AND PATHWAYS LEADING TO VARIOUS SPECIFIC END PRODUCTS.
CC -|- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U12012; AAA92668.1; -
DR EMBL: U12013; AAA92669.1; -
DR EMBL: U39404; AAB42382.1; -
DR EMBL: U39405; AAB42383.1; -
DR HSP: P08659; 1LC1.
DR InterPro: IPR000873; AMP-bind.
DR Pfam: PF00501; AMP-binding; 1.
DR PROSITE: PS00455; AMP_BINDING; 1.
KW Ligase; Phenylpropanoid metabolism.
FT VARIANT 422 422 E -> G.
SQ SEQUENCE 537 AA; 58590 MW; 6686EB2E84C7DC43 CRC64;

Query Match 27.5%; Score 773; DB 1; Length 537;
Best Local Similarity 35.1%; Pred. No. 2e-48;
Matches 187; Conservative 100; Mismatches 200; Indels 46; Gaps 10;

QY 26 YOSLYRFASFPEAIIIDAHTNEVISYQAIFETSCRLAVSIEQYGLNENNVVGCSENNINF 85
DB 32 FERVEAFADRP-CLIDGATDRTYCFSEVELISRKVAAGLAKGLQOGQVVMLLPNCIEF 90
QY 86 FNPVLAALYLGIPVATSNMYYTDGELTGHNLISKPTIMFSSKKALPLILRVQONLSFIKK 145
DB 91 AFVFMGASVRGAIVTANPFYKPE-----IAKQAKAGARIIVTAAAYVEKLADLQSH 144
QY 146 VVVIDSMYDINGVECVSTFVARYTDHTFDP-LSFTPKDFDPLEKIALIMSSSGTTGLPKG 204
DB 145 DVLVITIDDAKPEGQHISVLTAEATQCPAVKIHPDD-----VVALPYSGTGTGLPKG 198
QY 205 VVLSHSLTIR-----FVHSRDPYGTGRTVPQTSILSLVPFFHAFMGFTTLSY 252
DB 199 VMLTHGLVSSVAQVDGPNPNLYFHSDD-----VILCVLPFLHYLSNVLIC 247
QY 253 FV-VGLKVVMLKKEGALFKTIQNTIPTIVVAPPVWFLAKSPLVQDYDLSLSEVAT 311
DB 248 ALRAGAATLIMQKFNLTTCLELIQKYKVTVAPVPPVILDITKSPVQYDVSSVRIIMS 307
QY 312 GGAPLGKDAEVAEAKRLKPLGIIOGYGLTETCCAVMITPHNA-----VKTGSTGRPLPYI 366
```

```
Db 308 GAAPLGKELEDALRERFPAIFGQGYGHTACPVLANMLAFKNFPFVKSQSGGIVVRNA 367
QY 367 KAKVLDNATGKALGPGRGEICFCQSEMIMKGYNNPEATIDTDKGMHSGDIGYDED 426
Db 368 QIKILDTEGSLPHNQAGEICIRGPEIMKGYINDPESTAATIDEGLWHTGDVEYIDDD 427
QY 427 GNFFIVDRKLKELIKYKGYQVAPAELENLLLOHPSITADAGVTGVPDEFQGLPAACVVL 486
Db 428 EEIFIVDRVKEIKYKGFQVAPAELEALLVAHPSIADAAVVPQKHEEAGEVPAFVVKSS 487
QY 487 GKTLTEKEVQDFIAAQVTPTKHLRGVGVFVDSIPKGPCTGKLIRKELREIFAQR 539
Db 488 --EISEQEIKEFVAKQVIFYKKIH-RVYFVDAIPKSPSGKIILRKDLRSRLAAK 537
```

Search completed: September 4, 2002, 15:08:06  
Job time: 245 sec

**This Page Blank (uspio,**



Db 301 IDKYDLSNLHEITASGAPLSKEVGEAVAKRPHLPGRIOGYGTETTSAILITPEGDDKPG 360  
QY 358 STGRPLPYIKAKVLONATKALGPCGERGICQSEMIKMGYYNNPEATIDTDKDGWLHS 417  
Db 361 AVGVVPPFEAKVVDLTGKTLGVNQRGELCVRGPMISGYNNPEATNALIDKDGWLHS 420  
QY 418 GDIGYDEGNFFIVDRKLKELIKYGYQVAPAELENLLQHPISADAGVTGVPDEFEGQL 477  
Db 421 GDIAWDEHFFIVDRLSKLIKYGQVAPAELESILLQHPNIFDAGVAGLPDDAGEL 480  
QY 478 PAACVVLSESGKTLTEKEVQDTAAQVTPTKHLRGVGVFVDSIPKPGTKGLRKELREIF- 536  
Db 481 PAAVVLEHGKTMTEKEIVDVYASQVTTAKKLGRGVGVFVDEVPKGLTGKLDARKIREILI 540  
QY 537 -AORAPKSKL 545  
Db 541 KAKGKSKL 550  
RESULT 2  
S62787  
Photinus-luciferin 4-monooxygenase (ATP-hydrolyzing) (EC 1.13.12.7) - Lampyris noctiluca  
N:Alternate names: firefly-type luciferase  
C:Species: Lampyris noctiluca  
C:Date: 23-Aug-1996 #sequence\_revision 01-Nov-1996 #text\_change 26-May-2000  
R:Salal-Newby, G.B.; Thomson, C.M.; Campbell, A.K.  
Biochem. J. 313, 761-767, 1996  
A:Title: Sequence and biochemical similarities between the luciferases of the glow-worm  
A:Reference number: S62787; MUID:96190714  
A:Accession: S62787  
A:Molecule type: mRNA  
A:Residues: 1-547 <SAL>  
A:Cross-references: EMBL:X89479; NID:g899314; PIDN:CAA61668.1; PID:g899315  
C:Superfamily: 4-coumarate--CoA ligase; acetate--CoA ligase homology  
C:Keywords: ATP; luminescence; monooxygenase; oxidoreductase  
F:71-535/Domain: acetate--CoA ligase homology <ACL>  
Query Match 54.6%; Score 1534; DB 2; Length 547;  
Best Local Similarity 53.5%; Pred. No. 2e-101;  
Matches 293; Conservative 106; Mismatches 145; Indels 4; Gaps 3;  
QY 1 MEE-ENIRHGERPRDIVHPGSGAQOQLYSLYKFAFPE--AIIDAHTNEVISYAIQIFETS 57  
Db 1 MEDAKNIMHGPAPEYPLEDGTAGEQLHKAMKRYAQVPGTIAFTDAHAENVITYSEVFEMA 60  
QY 58 CRLAVSIEQYGLNENNVGVCSENNINFPNVLALYLGIPVATSNMYTDGELTCHLNI 117  
Db 61 CRLAETMKRYGLGLQHHIAVCSNLSQFFMPVCGALFTGVGVASTNDIYNRELYNSLSI 120  
QY 118 SKPTIMFSKKALPLTLRVQONLSIFIKVWVTDSDMYDINGVBCVSTFVARVTDHTFDPLS 177  
Db 121 SOPTIVSCSKRALQILGVQKLPILQIKIVILDSREDYMGKSMYSFTEHSLPAGFNEYD 180  
QY 178 FPPKDFPLEKIALIMSSSGTGLPKGVVLSHRSITIRFVHSRDIYGTRTVPQTSILSL 237  
Db 181 YIPDSFRETATALIMNSSGSLPKGVVLTQNVCFRSHCRDPVFGNQIIPDTAILTV 240  
QY 238 VPFHAFGCMFTLSFVVLGKLVMLKKFEGALFLKTIQNYKPIPTIVAPPVWVFLAKSPL 297  
Db 241 IPFHFGMFTTGLYTCGFRVLVRYFEELFSLRSLQDYKTSALLVPTLFSFFAKSTL 300  
QY 298 VDOYDLSLSEVATGAPLGKDVAEAVAKRLKPLGIIQCYGLTETCCAVMITPHNAVKTG 357  
Db 301 VDKYDLSNLHEITASGAPLAKKEVGEAVAKRFLKPIGRIOGYGLTETTSAILITPEGDDKPG 360  
QY 358 STGRPLPYIKAKVLONATKALGPCGERGICQSEMIKMGYYNNPEATIDTDKDGWLHS 417  
Db 361 ACQGVVPPFEAKIVDLDTGKTLGVNQRGELCVRGPMISGYNNPEATNALIDKDGWLHS 420  
QY 418 GDIGYDEGNFFIVDRKLKELIKYGYQVAPAELENLLQHPISADAGVTGVPDEFEGQL 477  
Db 481 PAAVVLEHGKTMTEKEIVDVYASQVTTAKKLGRGVGVFVDEVPKGLTGKLDARKIREILI 540

Db 421 GDIAWYDKGHFFIVDRLSKLIKYGQVAPAELESILLQHPFFIFDAGVAGIPDDAGEL 480  
QY 478 PAACVVLSESGKTLTEKEVQDTAAQVTPTKHLRGVGVFVDSIPKPGTKGLRKELREIFA 537  
Db 481 PAAVVLEHGKTMTEKEVQDYVAGVQVASKRLRGVGVFVDEVPKGLTGKLDARKIREIL- 539  
QY 538 ORAPKSKL 545  
Db 540 NMKGKSKL 547  
RESULT 3  
JS0181  
Photinus-luciferin 4-monooxygenase (ATP-hydrolyzing) (EC 1.13.12.7) - Genji firefly  
N:Alternate names: firefly luciferase  
C:Species: Luciola cruciata (Genji firefly)  
C:Date: 07-Jun-1990 #sequence\_revision 20-Aug-1994 #text\_change 26-May-2000  
C:Accession: JS0181  
R:Masuda, T.; Tatsumi, H.; Nakano, E.  
Gene 77, 265-270, 1989  
A:Title: Cloning and sequence analysis of cDNA for luciferase of a Japanese firefly,  
A:Reference number: JS0181; MUID:89326143  
A:Accession: JS0181  
A:Molecule type: mRNA  
A:Residues: 1-548 <MAS>  
A:Cross-references: GB:M26194; NID:g159050; PIDN:AAA29135.1; PID:g159051  
C:Comment: This protein catalyzes the oxidation of luciferin in the presence of ATP,  
pecies from green to yellow.  
C:Superfamily: 4-coumarate--CoA ligase; acetate--CoA ligase homology  
C:Keywords: ATP; luminescence; monooxygenase; oxidoreductase  
F:73-537/Domain: acetate--CoA ligase homology <ACL>  
Query Match 53.0%; Score 1491.5; DB 1; Length 548;  
Best Local Similarity 53.1%; Pred. No. 2.2e-98;  
Matches 289; Conservative 98; Mismatches 154; Indels 3; Gaps 2;  
QY 3 BENIRHGERPRDIVHPGSGAQOQLYSLYKFAFPE--AIIDAHTNEVISYAIQIFETSCRLA 61  
Db 7 DENIVVGPKPFPPIEGSGAGTQLRKYMERYAKLGAIAFTNAVTVGDVSYAEVLEKSCCLG 66  
QY 62 VSIEQYGLNENNVGVCSENNINFPNVLALYLGIPVATSNMYTDGELTCHLNIKPT 121  
Db 67 KALQVGLVVDGRIALCSENCPEFPPIVIAGLFIGVGVAPTNEIYTLRELVHSLGSKPT 126  
QY 122 IMFSSKKALPLTLRVQONLSIFIKVWVTDSDMYDINGVBCVSTFVARVTDHTFDPLSFTPK 181  
Db 127 IVFSSKKGLDKVITVQKTVTTIKTIVILDSKVYDYGVOCLDTFIRKNTPPGQASFKTV 186  
QY 182 DFDPLEKIALIMSSSGTGLPKGVVLSHRSITIRFVHSRDIYGTRTVPQTSILSLVFPFH 241  
Db 187 EVDREKQVALIMNSSSGTGLPKGVQVLTHTVTRFSHARDPIYGNQVSPGTAULTVVPFH 246  
QY 242 HAFGFTTILSFVVLGKLVMLKKFEGALFLKTIQNYKPIPTIVAPPVWVFLAKSLPVOY 301  
Db 247 HGFGMFTTGLYLCIGFRVVMVLTKFDEETFLKTLQDYKCTSVILVLTFLAILNKSELLKY 306  
QY 302 DLSLSEVATGAPLGKDVAEAVAKRLKPLGIIQCYGLTETCCAVMITPHNAVKTGSTGR 361  
Db 307 DLSNVEIASGAPLSKEVGEAVARRNLPGRVQGTGTETTSAILITPEGDDKPGASK 366  
QY 362 PLPTYIKAKVLONATKALGPCGERGICQSEMIKMGYYNNPEATIDTDKDGWLHSGDTG 421  
Db 367 VYPLFKAKVIDLDTKSLGPNRRGEVCKGPMKMGYYNNPEATKELIDEGMLHTGDIG 426  
QY 422 YDEGNEFFIVDRKLKELIKYGYQVAPAELENLLQHPISADAGVTGVPDEFEGQLPAAC 481  
Db 427 YYDEKHEFFIVDRLSKLIKYGQVAPAELESILLQHPISIFDAGVAGVDPVAGELPGAV 486  
QY 482 VYLESSKTLTEKEVQDTAAQVTPTKHLRGVGVFVDSIPKPGTKGLRKELREIFAQRAP 541  
Db 487 VYLESSKNTEKEVMDYVASQVSNKRLRGVGVFVDEVPKGLTGKLDARKIREIL--KKP 544



QY 542 KSKL 545

Db 545 VARL 548

## RESULT 4

S33788

Photinus-luciferin 4-monooxygenase (ATP-hydrolyzing) (EC 1.13.12.7) [validated] - southern  
N:Alternate names: firefly luciferase  
C:Species: Luciola mingrelia (southern Russian firefly)  
C:Date: 02-Dec-1993 #sequence\_revision 03-Nov-1995 #text\_change 02-Jun-2000  
C:Accession: S33788

R:Devine, J.H.; Kutuzova, G.D.; Green, V.A.; Ugarova, N.N.; Baldwin, T.O.

Biochim. Biophys. Acta 1173, 121-132, 1993  
A:Title: Luciferase from the East European firefly Luciola mingrelia: cloning and nucleotide sequence  
A:Reference number: S33788; MUID:93277944

A:Accession: S33788

A&gt;Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-548 &lt;DEV&gt;

A:Cross-references: GB:S61961; NID:9409316; PIDN:AAB26932.1; PID:9409317

C:Superfamily: 4-coumarate--CoA ligase; acetate--CoA ligase homology

C:Keywords: ATP; luminescence; monooxygenase; oxidoreductase

F:72-537/Domain: acetate--CoA ligase homology &lt;ACL&gt;

## Query Match

Best Local Similarity 52.7%; Score 1483; DB 2; Length 548;

Matches 284; Conservative 107; Mismatches 151; Indels 4; Gaps 3;

QY 2 EENIRHGERPRDVIHPGSGAQQOYLSLYKFASFPPE-AIIDAHTNEVISVAQIFETSCRL 60

Db 5 KEENVYGGPLPFYPIEGSGAQQLHKYMOYAKLGAIAFSNALTGVDSIQEYFDITCRL 64

QY 61 AVSYEGYGLNENNVGVCSENNINFPNVLAAALYLGPVATSDMYTDELGTGHLNISRP 120

Db 65 AZAMNFGKHPHEALCALSCENCEEFFIPVLGLYGVAVAPNEIYTLRELHNSLGIAP 124

QY 121 TMFSKKALPILRVQOONLSFKKVVVSDMYDINGVECVSTFVARYTDHFDPLSFTFP 180

Db 125 TVFSSKGLPKVLEQVTKVTCIKKIVILDSKVNFQGHDCMETFIKKHVELGFPSSFPV 184

QY 181 KDF-DPLEKIALIMSSGCTGLPKGVVLSHRSLSLIRFVHSRDPYIGTRVPOTSILSLVP 239

Db 185 IDVKNRKHQVALLMNSGSGTGLPKGVRLITHEGAVTRFSHAKDPIYGNQVSPGTAILTVVP 244

QY 240 FHAFGMFTTSLYFVVGKLVVLMKKFEGALFLKTIQNYKIPTIVVAPPVVMFLAKSPVLD 299

Db 245 FHHGFMFTTGLYFAGCYRVVMLTDFEELFLTLQDYKTSVILVPTLFAILNKSLELD 304

QY 300 QYDLSLLEVATGGAPLGDVAEAVAKRLKLPGLIIOGYLGTETCCAVMITPHNAVKTGST 359

Db 305 KFDLSNLTEIASGAPLAKVEGEAVARRNLPVGRQGYGLTETTSFAIITPEGDDKPGAS 364

QY 360 GRPLPIKAKVLDNATGKALGGERGEICFQSEMIMKGYNNPEATIDIDKQHLHSGD 419

Db 365 GKVPFLPKVVDLDTKTLGVNRRGEICVKSPLMGLYSNNPEATREIDEGLHTGD 424

QY 420 IGYDDEGNFFIVDRUKELIKYGYOVAPEALENLLQHPSTADAGVTGVPDEFGSLPA 479

Db 425 IGYDDEHFFIVDRUKSLIKYGYOVPAELESVLLQHPNIFDAGVAGVDPDAGELPG 484

QY 480 ACVVLESKLTLEKEVODFIAAQVTPTKHLRGVVFVDSIPKGPCTGKLIRKELREIFAOR 539

Db 485 AVVMEKGTMTKEIVDYVNSQVNVNKKLRGGVRFVDEVPKGLTCKIDAKVIREIL--K 542

QY 540 AKSKL 545

Db 543 KQAKM 548

## RESULT 5

S23437

Photinus-luciferin 4-monooxygenase (ATP-hydrolyzing) (EC 1.13.12.7) - Japanese firefly  
N:Alternate names: firefly luciferase  
C:Species: Luciola lateralis (Japanese firefly)  
C:Date: 22-Jan-1993 #sequence\_revision 20-Aug-1994 #text\_change 02-Jun-2000  
C:Accession: S23437; S57417

R:Tatsumi, H.; Kajiyama, N.; Nakano, E.

Biochim. Biophys. Acta 1131, 161-165, 1992

A:Title: Molecular cloning and expression in Escherichia coli of a cDNA clone encoding  
A:Reference number: S23437; MUID:92305054

A:Accession: S23437

A:Molecule type: mRNA

A:Residues: 1-548 &lt;TAT&gt;

A:Cross-references: EMBL:X66919; NID:95526; PIDN:CAAA7358.1; PID:95527

R:Cho, K.; Choi, Y.; Boo, K.

submitted to the EMBL Data Library, June 1995

A:Description: Molecular cloning of gene for luciferase in Luciola lateralis.

A:Reference number: S57417

A:Accession: S57417

A:Molecule type: DNA

A:Residues: 1-4,'D',6-14,'K',16-145,'N',147-175,'P',177-505,'G',507-548 &lt;CHO&gt;

A:Cross-references: EMBL:249891; NID:9871400; PIDN:CAA90072.1; PID:9871401

C:Genetics:

C:Function: 114/1; 225/2; 332/2; 450/3; 506/1

A:Note: catalyzes the irreversible reaction of luciferin with oxygen and ATP to

C:Superfamily: 4-coumarate--CoA ligase; acetate--CoA ligase homology

C:Keywords: ATP; luminescence; magnesium; monooxygenase; oxidoreductase; peroxisome

F:73-537/Domain: acetate--CoA ligase homology &lt;ACL&gt;

## Query Match

Best Local Similarity 51.9%; Score 1458.5; DB 1; Length 548;

Matches 282; Conservative 101; Mismatches 158; Indels 3; Gaps 2;

QY 3 EENIRHGERPRDVIHPGSGAQQOYLSLYKFASFPPE-AIIDAHTNEVISVAQIFETSCRL 61

Db 7 DENIVYGGPFPYPIEGSGAQQLRKYMRYAKLGAIAFTNALTGVDYTYAEYLEKSCCLG 66

QY 62 VSTEQYGLNENNVGVCSENNINFPNVLAAALYLGPVATSDMYTDELGTGHLNISKPT 121

Db 67 EALKNYGLVVDGRIALCSECEFFIPVLGLYGVAVAPNEIYTLRELHNSLGIKPT 126

QY 122 TMFSKKALPILRVQOONLSFKKVVVSDMYDINGVECVSTFVARYTDHFDPLSFTPK 181

Db 127 TVFSSKGLPKVLEQVTKVTAIKTIVILDSKVYRGYSMDNFIKNTPGQFGSSFTV 186

QY 182 DFDPLEKIALIMSSGCTGLPKGVVLSHRSLSLIRFVHSRDPYIGTRVPOTSILSLVPFH 241

Db 187 EYVNRKQVALLMNSGSGTGLPKGVQLTHENAVTRFSHARDPIYGNQVSPGTAILTVVPFH 246

QY 242 FHAFGMFTTSLYFVVGKLVVLMKKFEGALFLKTIQNYKIPTIVVAPPVVMFLAKSPVLD 301

Db 247 HFGFMFTTGLYTCGRFVYMLTKFDEETFLKTLQDYKCSVILVPTLFAILNSELDDKY 306

QY 302 DLSLLEVATGGAPLGDVAEAVAKRLKLPGLIIOGYLGTETCCAVMITPHNAVKTGST 361

Db 307 DLSNLVEIASGAPLSKEIGEAVARRNLPVGRQGYGLTETTSAAIITPEGDDKPGASGK 366

QY 362 PLPIYIKAKVLDNATGKALGGERGEICFQSEMIMKGYNNPEATIDIDKQHLHSGDIG 421

Db 367 VVPLFAKAKVIDLDTKTLGNRRGECVCKPMLMKGVDPNPEATREIIDEGLHTGDIG 426

QY 422 YDDEDNFFIVDRUKELIKYGYOVAPEALENLLQHPSTADAGVTGVPDEFGQLPAAC 481

Db 427 YDDEEHFFIVDRUKSLIKYGYOVPAELESVLLQHPNIFDAGVAGVDPDAGELPGAV 486

QY 482 VVLESKLTLEKEVODFIAAQVTPTKHLRGVVFVDSIPKGPCTGKLIRKELREIFAQAP 541

Db 487 VVLEKSKMTKEVMDYVASQVSNKRLRGVRFVDEVPKGLTGKIDKAIKIREIL--KKP 544

QY 542 KSKL 545

: : :

Db 545 VAKM 548

## RESULT 6

S29354

Photinus-luciferin 4-monooxygenase (ATP-hydrolyzing) (EC 1.13.12.7) YE [validated] - lum  
N:Alternate names: firefly-type luciferase  
C:Species: Pyrophorus plagiophthalmus  
C:Date: 22-Nov-1993 #sequence\_revision 01-Sep-1995 #text\_change 02-Jun-2000  
C:Accession: S29354

R:Wood, K.V.; Lam, Y.A.; Selliger, H.H.; McElroy, W.D.

Science 244, 700-702, 1989  
A:Title: Complementary DNA coding click beetle luciferases can elicit bioluminescence of  
A:Reference number: S29352; MUID:89242142

A:Accession: S29354

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-543 &lt;WOO&gt;

C:Superfamily: 4-coumarate--CoA ligase; acetate--CoA ligase homology

C:Keywords: ATP; luminescence; monooxygenase; oxidoreductase

F:70-532/Domain: acetate--CoA ligase homology &lt;ACL&gt;

Query Match 49.5%; Score 1393; DB 2; Length 543;

Best Local Similarity 48.9%; Pred. No. 2.2e-91;

Matches 264; Conservative 113; Mismatches 161; Indels 2; Gaps 1;

QY 3 EENIRHGERPRDIVHPGSAQOOLYQSLYKFASFPEALIDAHTNEVISYQAIFETSCRLAV 62

Db 5 ERNVIYGPPLPLEDKTAGEMLFRAIRKHSHPQAIQVDFGDESLSYKEFFPEATCLLAA 64

QY 63 STEQYGLNENNVVGVCSENNINFPNVLAAALYLGIPVATSDMYTDGELTGHNLISKPRTI 122

Db 65 SLHNCGYKNDVVSICAENKREFPIIAAWYIGMIVAPVNESYIPDELCKYMGISKPQI 124

QY 123 MFSSKKALPLILRVQONLSFIKKVVVIDMYDINGVECVSTFVARYTDHTFDPLSTPKD 182

Db 125 VFCTKNILNKVLEVSQRTNFIKRIIILDTVENIHGCESLPNFISRYSDGNI--ANFKPLH 182

QY 183 FQPLEKIALIMSSSGTGLPKGVVLSHRSLTIRFVHSRDPITVYGRVTPVOTSLSLVPPFH 242

Db 183 YDPVEQVAAILCSSGTGLPKGMQTHQICVRLIHALDPEAGTQIPGVTVLVVYPPFH 242

QY 243 AFGMFTTSLYFVVGKLVMLKFEFALFKTIQNYKIPTIVVAPVPMVFLAKSPLVDQYD 302

Db 243 AFGFSLNLYFVVGKLVMLRREFEQAFLKAIQDYEVRSIVNVVPAIILFLSKSPLVDKYD 302

QY 303 LSLSTEVAATGGAPLKDVAEAVAKRLKLPGLIIOGYGLTETCCAVMTTPHNAVKTGSTRP 362

Db 303 LSSLRELCCGAAPLAKVEAEIAVKRLNLPGLRCGFGLTSTSANIHSLGDEFKSGSLGRV 362

QY 363 LPYIKAKVLNDNATGKALGPGEIGEICFQSEMIMKGYNNPEATIDTIDKGMVLSHGDIGY 422

Db 363 TPLMAAKIADRTGKALGPNGVGEICIKGPMVSKGVNNVNEATKEAIDDDGMVLSHGDIGY 422

QY 423 YDEGNGFFIVDLRLKELIKYQVAPAELENLLQHPISADAGVTGVPDFGGQLPAACV 482

Db 423 YDEDEHYVVDYRKELIKYKGSQVAPAELEILLNKPICRDAVAVGIPDLEAGELPSAFV 482

QY 483 VLESKTLTEKEVQDPIAAQVTPTKHLRGVGVFVDSIPKGTGKLIKRELREIFAQRAK 542

Db 483 VIQPGKEITAKEVYDYLAEVRSHTKYLRGVRFVDSIPRNVTKITRKELLKOLLEKSSK 542

## RESULT 7

S29355

Photinus-luciferin 4-monooxygenase (ATP-hydrolyzing) (EC 1.13.12.7) OR [validated] - lum  
N:Alternate names: firefly-type luciferase  
C:Species: Pyrophorus plagiophthalmus  
C:Date: 22-Nov-1993 #sequence\_revision 03-Nov-1995 #text\_change 02-Jun-2000  
C:Accession: S29355

R:Wood, K.V.; Lam, Y.A.; Selliger, H.H.; McElroy, W.D.

Science 244, 700-702, 1989

A:Title: Complementary DNA coding click beetle luciferases can elicit bioluminescence  
A:Reference number: S29352; MUID:89242142

A:Accession: S29355

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-543 &lt;WOO&gt;

C:Superfamily: 4-coumarate--CoA ligase; acetate--CoA ligase homology

C:Keywords: ATP; luminescence; monooxygenase; oxidoreductase

F:70-532/Domain: acetate--CoA ligase homology &lt;ACL&gt;

Query Match 49.4%; Score 1388; DB 2; Length 543;

Best Local Similarity 48.7%; Pred. No. 5.1e-91;

Matches 263; Conservative 114; Mismatches 161; Indels 2; Gaps 1;

QY 3 EENIRHGERPRDIVHPGSAQOOLYQSLYKFASFPEALIDAHTNEVISYQAIFETSCRLAV 62

Db 5 ERNVIYGPPLPLEDKTAGEMLFRAIRKHSHPQAIQVDFGDESLSYKEFFPEATCLLAA 64

QY 63 STEQYGLNENNVVGVCSENNINFPNVLAAALYLGIPVATSDMYTDGELTGHNLISKPRTI 122

Db 65 SLHNCGYKNDVVSICAENKREFPIIAAWYIGMIVAPVNESYIPDELCKYMGISKPQI 124

QY 123 MFSSKKALPLILRVQONLSFIKKVVVIDMYDINGVECVSTFVARYTDHTFDPLSTPKD 182

Db 125 VFCTKNILNKVLEVSQRTNFIKRIIILDTVENIHGCESLPNFISRYSDGNI--ANFKPLH 182

QY 183 FQPLEKIALIMSSSGTGLPKGVVLSHRSLTIRFVHSRDPITVYGRVTPVOTSLSLVPPFH 242

Db 183 YDPVEQVAAILCSSGTGLPKGMQTHQICVRLIHALDPEAGTQIPGVTVLVVYPPFH 242

QY 243 AFGMFTTSLYFVVGKLVMLKFEFALFKTIQNYKIPTIVVAPVPMVFLAKSPLVDQYD 302

Db 243 AFGFSLNLYFVVGKLVMLRREFEQAFLKAIQDYEVRSIVNVVPAIILFLSKSPLVDKYD 302

QY 303 LSLSTEVAATGGAPLKDVAEAVAKRLKLPGLIIOGYGLTETCCAVMTTPHNAVKTGSTRP 362

Db 303 LSSLRELCCGAAPLAKVEAEIAVKRLNLPGLRCGFGLTSTSANIHSLGDEFKSGSLGRV 362

QY 363 LPYIKAKVLNDNATGKALGPGEIGEICFQSEMIMKGYNNPEATIDTIDKGMVLSHGDIGY 422

Db 363 TPLMAAKIADRTGKALGPNGVGEICIKGPMVSKGVNNVNEATKEAIDDDGMVLSHGDIGY 422

QY 423 YDEGNGFFIVDLRLKELIKYQVAPAELENLLQHPISADAGVTGVPDFGGQLPAACV 482

Db 423 YDEDEHYVVDYRKELIKYKGSQVAPAELEILLNKPICRDAVAVGIPDLEAGELPSAFV 482

QY 483 VLESKTLTEKEVQDPIAAQVTPTKHLRGVGVFVDSIPKGTGKLIKRELREIFAQRAK 542

Db 483 VIQPGKEITAKEVYDYLAEVRSHTKYLRGVRFVDSIPRNVTKITRKELLKOLLEKSSK 542

## RESULT 8

S29352

Photinus-luciferin 4-monooxygenase (ATP-hydrolyzing) (EC 1.13.12.7) GR [validated] -  
N:Alternate names: firefly-type luciferase  
C:Species: Pyrophorus plagiophthalmus

C:Date: 22-Nov-1993 #sequence\_revision 01-Sep-1995 #text\_change 02-Jun-2000

C:Accession: S29352

R:Wood, K.V.; Lam, Y.A.; Selliger, H.H.; McElroy, W.D.

Science 244, 700-702, 1989

A:Title: Complementary DNA coding click beetle luciferases can elicit bioluminescence

A:Reference number: S29352; MUID:89242142

A:Accession: S29352

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-543 &lt;WOO&gt;

C:Superfamily: 4-coumarate--CoA ligase; acetate--CoA ligase homology

C:Keywords: ATP; luminescence; monooxygenase; oxidoreductase

F:70-532/Domain: acetate--CoA ligase homology &lt;ACL&gt;

Query Match

49.1%; Score 1382; DB 2; Length 543;

```
Best Local Similarity 48.0%; Pred. No. 1.4e-90;
Matches 259; Conservative 118; Mismatches 161; Indels 1; Gaps 1;

Qy 3 ENTRHGRPRDIVHPGSAGQOOLYOSLYKFASFPAIRIDAHTNEVISYAQIFETSCRILAV 62
   I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I:::
Db 5 EKNVYIGPEPLHPLDLTAGEMLFRALKRHSHPALQALVDVYGEEMISYKEFFETCLLAQ 64
   I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I:::
Qy 63 STEQYGLNENNVGVCSENNINFFNPVLAALYLGTFPVATSDNMDYDGLTGHNLTSKPTI 122
   I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I:::
Db 65 SLHNCGYKMSDVISCAENNRFFPIIAAAYIGMIVAPVNEGYPDELCKVMGISRPOL 124
   I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I:::
Qy 123 MFSSKALPLILRVQOONLSFIKKVVVIDSMYDINGVECVSTFVARYTDHTDPLSFTPKD 182
   I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I:::
Db 125 VFCTRNILNKVLEQSRDFTIKRIIILDAVENIHGCSLNPFIISYSDGNI--ANFKPLH 182
   I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I:::
Qy 183 FDPLEKIALINSSSTGTTGLPKGVLSHRSLTIRFVHSDRPIYGTTRVTPOTSILSLVPHHH 242
   I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I:::
Db 183 YDPVEQVAAILCSSTGTTGLPKGMVTHQNICVRLIHALDPRAGTQOLIPGTVLVLYLPFFH 242
   I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I:::
Qy 243 AFGMPTTISYFVVGKLVNMLKFFEGALFLKTIQNYKIPTIVVAPVPMVFLAKSPLDVQYD 302
   I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I:::
Db 243 AFGFSINLGYFMVGLURVIMLRRFDQEAFLKAIQDYEVRSVINVPAILIILFLSKSPLVDKYD 302
   I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I:::
Qy 303 LSSLETEVATGGAPLKGKVAEAVAKRLKLPGLIITQGYGLTETCCAVMITPHNAVKTSTGRP 362
   I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I:::
Db 303 LSSLELCCGAAPLAKEVAEAVAKRLNLPGLIRCGFGLTESTSANIHSLCDDEKFSGLGRV 362
   I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I:::
Qy 363 LPYIKAKVLDNATGKALGGERGEICFQSEMIMKGYNNPEATIDTIDKDGWLHSGDIGY 422
   I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I:::
Db 363 TPLMAAKIADRETGKALGNQVGEICVKPMVSKGVNNVEATKEAIDDDGWLHSGDFGY 422
   I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I:::
Qy 423 YDEGNFFIVDLRLKELIKYGVQVAPAELENLLLOHPSIADAGVTVGPDDEFGGOLPAACV 482
   I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I:::
Db 423 AFGFSINLGYFMVGLURVIMLRRFDQEAFLKAIQDYEVRSVINVPAILIILFLSKSPLVDKYD 302
   I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I:::
Qy 303 LSSLETEVATGGAPLKGKVAEAVAKRLKLPGLIITQGYGLTETCCAVMITPHNAVKTSTGRP 362
   I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I:::
Db 303 LSSLELCCGAAPLAKEVAEAVAKRLNLPGLIRCGFGLTESTSANIHSLCDDEKFSGLGRV 362
   I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I:::
Qy 363 LPYIKAKVLDNATGKALGGERGEICFQSEMIMKGYNNPEATIDTIDKDGWLHSGDIGY 422
   I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I:::
Db 363 TPLMAAKIADRETGKALGNQVGEICVKPMVSKGVNNVEATKEAIDDDGWLHSGDFGY 422
   I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I:::
Qy 423 YDEGNFFIVDLRLKELIKYGVQVAPAELENLLLOHPSIADAGVTVGPDDEFGGOLPAACV 482
   I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I:::
Db 423 YDEGNFFIVDLRLKELIKYGVQVAPAELENLLLOHPSIADAGVTVGPDDEFGGOLPAACV 482
   I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I:::
Qy 483 VLESKTLTEKEVQDIFAAQVTPTKHLRGVVFVDSIPKPGTKLIRKELREIFAQAPK 542
   I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I:::
Db 483 VIQPKETAKEVYDYLAEVSHTYLKGVRVFDSDIPRNVGTGKITRKLKQLLEKSSK 542
   I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I:::
RESULT 9
Photinus-luciferin 4-monooxygenase (ATP-hydrolyzing) (EC 1.13.12.7) YG [validated] - lum
A:Alternate names: firefly-type luciferase
C:Species: Pyrophorus plagiophthalmus
C>Date: 25-Feb-1994 #sequence_revision 01-Sep-1995 #text_change 02-Jun-2000
C:Accession: S29353
R:Wood, K.V.; Lam, Y.A.; Seliger, H.H.; McElroy, W.D.
Science 244, 700-702, 1989
A:Title: Complementary DNA coding click beetle luciferases can elicit bioluminescence of
A:Reference number: S29352; MUID:89242142
A:Accession: S29353
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-543 <WOO>
C:Superfamily: 4-coumarate--CoA ligase; acetate--CoA ligase homology
C:Keywords: ATP; luminescence; monooxygenase; oxidoreductase
F:70-532/Domain: acetate--CoA ligase homology <ACL>

Query Match 49.0%; Score 1377; DB 2; Length 543;
Best Local Similarity 48.0%; Pred. No. 3.1e-90;
Matches 259; Conservative 117; Mismatches 162; Indels 1; Gaps 1;

Qy 3 ENTRHGRPRDIVHPGSAGQOOLYOSLYKFASFPAIRIDAHTNEVISYAQIFETSCRILAV 62
   I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I:::
Db 5 EKNVYIGPEPLHPLDLTAGEMLFRALKRHSHPALQALVDVYGEEMISYKEFFETCLLAQ 64
   I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I:::
Qy 63 STEQYGLNENNVGVCSENNINFFNPVLAALYLGTFPVATSDNMDYDGLTGHNLTSKPTI 122
   I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I:::
Db 65 SLHNCGYKMSDVISCAENNRFFPIIAAAYIGMIVAPVNEGYPDELCKVMGISRKQI 124
   I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I:::
Qy 123 MFSSKALPLILRVQOONLSFIKKVVVIDSMYDINGVECVSTFVARYTDHTDPLSFTPKD 182
```

```
Db 125 VFCTRNILNKVLEQSRDFTIKRIIILQTVENIHGCSLNPFIISYSDGNI--ANFKPLH 182
   I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I:::
Qy 183 FDPLEKIALINSSSTGTTGLPKGVLSHRSLTIRFVHSDRPIYGTTRVTPOTSILSLVPHHH 242
   I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I:::
Db 183 YDPVEQVAAILCSSTGTTGLPKGMVTHQNICVRLIHALDPRAGTQOLIPGTVLVLYLPFFH 242
   I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I:::
Qy 243 AFGMPTTISYFVVGKLVNMLKFFEGALFLKTIQNYKIPTIVVAPVPMVFLAKSPLDVQYD 302
   I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I:::
Db 243 AFGFSINLGYFMVGLURVIMLRRFDQEAFLKAIQDYEVRSVINVPAILIILFLSKSPLVDKYD 302
   I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I:::
Qy 303 LSSLETEVATGGAPLKGKVAEAVAKRLKLPGLIITQGYGLTETCCAVMITPHNAVKTSTGRP 362
   I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I:::
Db 303 LSSLELCCGAAPLAKEVAEAVAKRLNLPGLIRCGFGLTESTSANIHSLCDDEKFSGLGRV 362
   I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I:::
Qy 363 LPYIKAKVLDNATGKALGGERGEICFQSEMIMKGYNNPEATIDTIDKDGWLHSGDIGY 422
   I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I:::
Db 363 TPLMAAKIADRETGKALGNQVGEICVKPMVSKGVNNVEATKEAIDDDGWLHSGDFGY 422
   I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I:::
Qy 423 YDEGNFFIVDLRLKELIKYGVQVAPAELENLLLOHPSIADAGVTVGPDDEFGGOLPAACV 482
   I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I:::
Db 423 YDEGNFFIVDLRLKELIKYGVQVAPAELENLLLOHPSIADAGVTVGPDDEFGGOLPAACV 482
   I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I:::
Qy 483 VLESKTLTEKEVQDIFAAQVTPTKHLRGVVFVDSIPKPGTKLIRKELREIFAQAPK 542
   I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I:::
Db 483 VKQPKETAKEVYDYLAEVSHTYLKGVRVFDSDIPRNVGTGKITRKLKQLLEKSSK 542
   I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I:::
RESULT 10
4-coumarate--CoA ligase (EC 6.2.1.12) - common tobacco
A:Alternate names: 4-coumaroyl-CoA synthetase
C:Species: Nicotiana tabacum (common tobacco)
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000
C:Accession: T02074
R:Katayama, Y.; Kawai, S.; Morohoshi, N.; Kajita, S.
submitted to the EMBL Data Library, December 1994
A:Description: Cloning and nucleotide sequence of 4-coumarate:coenzyme A ligase gene
A:Reference number: Z14545
A:Accession: T02074
A:Status: preliminary; translated from GB/EMBL/DBDJB
A:Molecule type: mRNA
A:Residues: 1-542 <KAT>
A:Cross-references: EMBL:D43773; PIDN:BAA07828.1
C:Superfamily: 4-coumarate--CoA ligase; acetate--CoA ligase homology
C:Keywords: acid-thiol ligase; coenzyme A; flavonoid biosynthesis
F:73-532/Domain: acetate--CoA ligase homology <ACL>

Query Match 30.7%; Score 864.5; DB 1; Length 542;
Best Local Similarity 35.6%; Pred. No. 9.8e-54;
Matches 203; Conservative 112; Mismatches 194; Indels 61; Gaps 16;

Qy 3 ENIRHGE-----RPRDIVHPGSAGQOOLY--OSLYKFASFPAIRIDAHTNEVISYAQIFE 55
   I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I:::
Db 2 EKDTKHGDIIFRSKLPDIYIPNHLPLHSVCFENISEFSSRP-CLINGANKOYITVADVEL 60
   I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I:::
Qy 56 TSCRILAVSTEQYGLNENNVGVCSENNINFFNPVLAALYLGIPVATSDNMDYDGLTGHHL 115
   I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I:::
Db 61 SSRKVAAGLHKOGIOOKDTIMILLPNSPEVFAFAGASYLGAISTMANPLFTAEEVKQV 120
   I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I:::
Qy 116 NTSKRTIMFSSKKALPLI--LRVQONLSFIKKVVVIDSMYDINGVECVSTFVARYTD-HT 172
   I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I:::
Db 121 KASGAKIIVTQACHNVKVDYALENNV-----KIICIDSAPE----GCLHFSVLTOADEHD 172
   I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I:::
Qy 173 FDPLSFTPKDPLDKLEKIALIMSSSGTTGLPKGVLSHRSLTIR-----FVHRS 220
   I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I:::
Db 173 IPEVEIQPD-----VVALPYSSSGTTGLPKGVMLTHKGLVTSVAQOQVNGENRNLIHSE 226
   I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I:::
Qy 221 DPIYGTTRVTPOTSILSLVPHHAFGMFTTSLYFV-VGLKVMMLKPKFEGALFLKTIQNYKI 279
   I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I::: I:::
Db 227 D-----VLLCVLPPLPHIYSLNSVLLCGLRGAAILIMQKFDIVPFLLEIQNYKV 275
```

Qy 280 PTIVVAPVMVFLAKSPLVDDYDSSLSLEVATGGAPLKDAEAVAKRLKLP--IIQGY 337  
Db 276 TIGPVPPIVLAIAKSPWDDYDSSSVTVMSGAAPLKELEDTV--RAKFPNAKLGQY 333  
Qy 338 GLTET-----CCAVMITPHNAVKTGSRPLPYTKAKVLONATGKALGPGERGEICFQS 391  
Db 334 GMTAGPVLNCLAFAPKPEF--IKSGACGTVVRNAEMKIVDPETGNSLPRNQSGEICIRG 392  
Qy 392 EMIMKGYNNPEATIDTIDKDGWLHSGDIGYVVDGNGFFIVDRKLKELIKYGYOVAPAE 451  
Db 393 DQIMKGYLNDEATARTIDKEGWLYTGIDIGYIDDDDELFIIVDRKLKELIKYGFQVAPAE 452  
Qy 452 ENLLQHPISADAGVTGVPDEFGGOLPAACVVLSEKTLTEKEVQDFIAAQTPTPKHLRG 511  
Db 453 EALLNHPFTSDAAVPMKDEQAEEVPAFVVRSSGSTITEDEVKDFISKQVIFYKRIK- 511  
Qy 512 GVVFVDSPTKPGTKLIRKELREIFAQAP 541  
Db 512 RVFFVDAVPKSPSGKILRKDLRAKLAAGLP 541

RESULT 11  
T03789  
4-coumarate--CoA ligase (EC 6.2.1.12) 4CL2 - common tobacco  
N:Alternate names: 4-coumaroyl-CoA synthetase  
C:Species: Nicotiana tabacum (common tobacco)  
C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 11-May-2000  
C:Accession: T03789  
R:Lee, D.; Douglas, C.J.  
Plant Physiol. 112, 193-205, 1996  
A:Title: Two divergent members of a tobacco 4-coumarate:coenzyme A ligase (4CL) gene fam  
A:Reference number: Z15086; MUID:96416441  
A:Accession: T03789  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-542 <LEE>  
A:Cross-references: EMBL:U50846; NID:g1663723; PIDN:AAB18638.1; PID:g1663724  
C:Genetics:  
A:Gene: 4CL2  
C:Superfamily: 4-coumarate--CoA ligase; acetate--CoA ligase homology  
C:Keywords: acid-thiol ligase; coenzyme A; flavonoid biosynthesis  
F:73-532/Domain: acetate--CoA ligase homology <ACL>

Query Match 30.4%; Score 855; DB 2; Length 542;  
Best Local Similarity 36.1%; Pred. No. 4.7e-53;  
Matches 195; Conservative 109; Mismatches 182; Indels 54; Gaps 14;

Qy 26 YQSLYKFAFPFAIIDAHTNEVISAQIFETSCRVAISIEQYGLNENNVGVCSENNIN 85  
Db 32 FENISEFSRPP-CLINGANKQIITYADVELNSRKVAAGLHKQIGQPKDTIMILLPNSPEF 90  
Qy 86 FNPVLAALYLIPVATSDMTDGLTGLHLMISKPTIMFSSKKALPLILRVQOONSFTK- 144  
Db 91 VFAFTGASYLGAISTMANPLETPAEVVKQAKASSAKIVTQACHNVK----KQYAFEND 146  
Qy 145 -KVYVIDSMYDINGVECVS-TEVARYTDHTFDPLSFTPKDPPLEKIALINSSSGTTGLP 202  
Db 147 VKIICIDSAPE---GCLHFSLTQANHEHDIPEVQIPDD-----VVALPYSSGTTGLP 196  
Qy 203 KGVLSHRLSLTIR-----FVHSRDPITYGTRTPOTSILSLVPFFHAFGMFTTL 250  
Db 197 KGVMLTKGLVTSVAQVVDGENPNLYIHSD-----VMLCVLPPLFIHYSLSVL 245  
Qy 251 SYFV-VGLKVVMLKKEGALFLKTIQNYKIPTIVVAPPVWFLAKSPLVDDYDSSLSLEV 309  
Db 246 LCGRVGAAILIMQKFDIVSLELTQRYKVTIGPVPPIVLAIAKSPWDDYDSSSVRTV 305  
Qy 310 ATGGAPLGKDAEAVAKRLKLP--IIQGYGLTET-----CCAVMITPHNAVKTGSR 361  
Db 306 MSGAAPLKELEDTV--RAKFPNAKLGQGYGTWTEAGPVLNCLAFAPKPEF--IKSGACGT 362  
Qy 362 PLPIYAKVLONATGKALGPGERGEICFQSEMINKGYNNPEATIDTIDKDGWLHSGDIG 421

Db 363 VVRNAEMKIVDPKTNLSLPRNQSGEICIRGQIMKGYLNDEATARTIDKEGWLYTGDIG 422  
Qy 422 YYDEGNGFFIVDRKLKELIKYGYOVAPAELENLLQHPISADAGVTGVPDEFGGOLPAAC 481  
Db 423 YIDDDDELFIIVDRKLKELIKYGFQVAPAELEALLNHPNISDAAVPMKDEQAEEVPAF 482  
Qy 482 VVLESKTLTEKEVQDFIAAQTPTPKHLRGVVFVDSIPKPGTKLIRKELREIFAQAP 541  
Db 483 VVRNSGSTITEDEVKDFISKQVIFYKRIK-RVFFVDAIPKSPSGKILRKDLRAKLAAGLP 541

RESULT 12  
A39827  
4-coumarate--CoA ligase (EC 6.2.1.12) 1 - potato  
N:Alternate names: 4-coumaroyl-CoA synthetase  
C:Species: Solanum tuberosum (potato)  
C:Date: 20-Mar-1992 #sequence\_revision 20-Aug-1994 #text\_change 05-May-2000  
C:Accession: A39827  
R:Becker-Andre, M.; Schulze-Lefert, P.; Hahlbrock, K.  
J. Biol. Chem. 266, 8551-8559, 1991  
A:Title: Structural comparison, modes of expression, and putative cis-acting elements  
A:Reference number: A39827; MUID:91217100  
A:Accession: A39827  
A:Molecule type: DNA  
A:Residues: 1-545 <BEC>  
A:Cross-references: GB:W62755; NID:g169573; PIDN:AAA33842.1; PID:g169574  
C:Superfamily: 4-coumarate--CoA ligase; acetate--CoA ligase homology  
C:Keywords: acid-thiol ligase; coenzyme A; flavonoid biosynthesis  
F:76-535/Domain: acetate--CoA ligase homology <ACL>

Query Match 30.1%; Score 847; DB 1; Length 545;  
Best Local Similarity 36.6%; Pred. No. 1.7e-52;  
Matches 196; Conservative 110; Mismatches 169; Indels 60; Gaps 15;

Qy 26 YQSLYKFAFPFAIIDAHTNEVISAQIFETSCRVAISIEQYGLNENNVGVCSENNIN 85  
Db 35 FENISEFSRPP-CLIDGANDRIYTYAEVELTSRKVAVGLNKLGIQOKDTIMILLPNCPEF 93  
Qy 86 FNPVLAALYLIPVATSDMTDGLTGLHLMISKPTIM-----FSSKKALPLILRVQOONL 140  
Db 94 VFAFTGASYLGAISTMANPLETPAEVVKQAKASSAKIVTQACFAGVK----DYAIENDL 150  
Qy 141 SFIKKVVVIDSMYDINGVECVS-TEVARYTDHTFDPLSFTPKDPPLEKIALINSSSGTT 199  
Db 151 ----KVICVDSVPE---GCVHFSLEIQSDHEIPDKVIQIPDD-----VVALPYSSGTT 196  
Qy 200 GLPKGVLSHRLSLTIR-----FVHSRDPITYGTRTPOTSILSLVPFFHAFGMF 247  
Db 197 GLPKGVMLTKGLVTSVAQVVDGENANLYMHSD-----VLMCVLPPLFIHYSLN 245  
Qy 248 TSLSYFV-VGLKVVMLKKEGALFLKTIQNYKIPTIVVAPPVWFLAKSPLVDDYDSSLS 306  
Db 246 SVLLCALRVGAAILIMQKFDIAQFLLEIPKHKVTIGPVPPIVLAIAKSPLVDDYDSSV 305  
Qy 307 TEVATGGAPLGKDAEAVAKRLKLP--IIQGYGLTET-----CCAVMITPHNAVKTGS 358  
Db 306 RTVMMSGAAPLKELEDV--RAKFPNAKLGQGYGTWTEAGPVLNCLAFAPKPEF--IKSGA 362  
Qy 359 TGRPLPYIKAKVLONATGKALGPGERGEICFQSEMINKGYNNPEATIDTIDKDGWLHSG 418  
Db 363 CGTVVRNAEMKIVDPDTCGLSPRNPQGEICIRGQIMKGYLNDEATARTIEKGLWHTG 422  
Qy 419 DIGYVDEGNGFFIVDRKLKELIKYGYOVAPAELENLLQHPISADAGVTGVPDEFGGOLP 478  
Db 423 DIGFIDDDDELFIIVDRKLKELIKYGFQVAPAELEALLNHPNISDAAVPMIDEGAVEP 482  
Qy 479 AACVVLSESKTLTEKEVQDFIAAQTPTPKHLRGVVFVDSIPKPGTKLIRKELR 533  
Db 483 VAFVVRNSGSTITEDEVKDFISKQVIFYKRIK-RVFFVETVPKSPSGKILRKDLR 536

## RESULT 13

T07908  
4-coumarate--CoA ligase (EC 6.2.1.12) 2 - western balsam poplar x cottonwood  
C:Species: Populus trichocarpa x Populus deltoides (western balsam poplar x cottonwood)  
C:Date: 14-May-1999 #sequence\_revision 14-May-1999 #text\_change 11-May-2000  
C:Accession: T07908  
R:Allina, S.M.; Pri-Hadash, A.; Theilmann, D.A.; Ellis, B.E.; Douglas, C.J.  
Plant Physiol. 116, 743-754, 1998  
A:Title: 4-Coumarate:Coenzyme A ligase in hybrid poplar. Properties of native enzymes, c  
A:Reference number: Z16208; MUID:98150279  
A:Accession: T07908  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: mRNA  
A:Residues: 1-548 <ALL>  
A:Cross-references: EMBL:AF008183; NID:g2911796; PIDN:AAC39365.1; PID:g2911797  
C:Genetics:  
A:Gene: 4CL2  
C:Superfamily: 4-coumarate--CoA ligase; acetate--CoA ligase homology  
C:Keywords: acid-thiol ligase; coenzyme A; flavonoid biosynthesis  
F:74-533/Domain: acetate--CoA ligase homology <ACL>

Query Match 30.1%; Score 847; DB 2; Length 548;  
Best Local Similarity 36.8%; Pred. No. 1.8e-52;  
Matches 196; Conservative 104; Mismatches 177; Indels 56; Gaps 15;

QY 26 YQSLYKFPPEALIDAHTNEVISAQIFETSCRLAVSIEQYGLNENNVGVCSENNINF 85  
DB 33 FEKLSQKNDP-CLINGPTGDIYADVELTSRVSAGLYKLQOQGVILLQLQNSPEF 91  
QY 86 FNPVLAALYLGPVATSDMTDGTGELTGLNLSKPTIMFSSKKKALPLTLRVOONLSFIKK 145  
DB 92 VFAFLGASFGAISTANPFTTSIAKQATASAKLIITHAAVAKVOQPAQENDHV-K 150  
QY 146 VVVIDMYDINGVECVSTFVARYDHTFDLP--LSFTPKDFDPLEKIALIMSSSGTGLPKG 204  
DB 151 IMTIDSLTE---NCLHFSELTSSENEIPTVKIKPDD-----IMALPYSSGTTGLPKG 200  
QY 205 VVLSHRSLSLIR-----FVHSRDPYIGTRTPVQTSILSLVPPHFAFGMFTLSY 252  
DB 201 VMLTHKGLVTSVAQQVDGNGENLNFYHERD-----VILCVPLFH---IYLSNV 246  
QY 253 FVVGK----VVMLKFFGALFLKTIQNYKIPTIVVAPVPMVFLAKSLVDQYDLSLITE 308  
DB 247 FLCGLRAGSAILVMKQFDVSLMDLVQYKVTIAPLVPPICIAKSPVVDQYDLSLIRT 306  
QY 309 VATGAPLKGKDAVAAKRLKLPG--IIQYGLTET-----CCAVMITPHNAVKTSTG 360  
DB 307 VLSGAAPLKEEDTV--RAKLNKLGQGYGTEAGPVIAMCLAFAPKEPE-IKSGAG 363  
QY 361 RPLPYIKAKVLNDATGKALGPGERGEICFQSEMIMKGYNNPEATIDTIDKGLHSGDI 420  
DB 364 TVWRNAEMKIIVDPETGESOPRNTKTEICIRGQIMKGYLNDPEATERTIDKGLWLTGDI 423  
QY 421 GYDDEGNFFIVDRKLKELIKYGVQVAPAELENLLQHPSTADAGVTCVPDEFGQLPAA 480  
DB 424 GYDDED-ELFTVDRKLKELIKYGVQVAPAELEAMLIAPHNDSDAAVPMKDEAAGEVPA 482  
QY 481 CWVLESKTLTEKEVQDFIAAQVPTPKHLRGVVFVDSIPKGTGKLIRKELR 533  
DB 483 FVVRNSGSKTTEDEIKQYISKQVIFYKRI-GRVFFTEIPRAPSGKILRKDLR 534

## RESULT 14

B39827  
4-coumarate--CoA ligase (EC 6.2.1.12) 2a - potato  
N:Alternate names: 4-coumaroyl-CoA synthetase  
C:Species: Solanum tuberosum (potato)  
C:Date: 20-Mar-1992 #sequence\_revision 20-Aug-1994 #text\_change 05-May-2000  
C:Accession: B39827  
R:Becker-Andre, M.; Schulze-Lefert, P.; Hahlbrock, K.  
J. Biol. Chem. 266, 8551-8559, 1991  
A:Title: Structural comparison, modes of expression, and putative cis-acting elements of

A:Reference number: A39827; MUID:91217100

A:Accession: B39827  
A:Molecule type: DNA  
C:Residues: 1-545 <BNC>  
C:Superfamily: 4-coumarate--CoA ligase; acetate--CoA ligase homology  
C:Keywords: acid-thiol ligase; coenzyme A; flavonoid biosynthesis  
F:76-535/Domain: acetate--CoA ligase homology <ACL>

Query Match 29.8%; Score 838; DB 1; Length 545;  
Best Local Similarity 36.4%; Pred. No. 7.6e-52;  
Matches 195; Conservative 109; Mismatches 171; Indels 60; Gaps 15;

QY 26 YQSLYKFPPEALIDAHTNEVISAQIFETSCRLAVSIEQYGLNENNVGVCSENNINF 85  
DB 35 FENLSEFNSRP-CLIDGANDRIYTAEVELTSRVAVGLNKLGIQOQDTIMILLPNCPEF 93  
QY 86 FNPVLAALYLGPVATSDMTDGTGELTGLNLSKPTIM-----FSSKKALPLTLRVOONL 140  
DB 94 VFAFIGASYLGAISTMANPLFTPAEVVQAKASSAKIVITOCAPGAKVK---DYAIENDL 150  
QY 141 SFIKKVVVIDMYDINGVECVS-TFVARYDHTFDPLSFTPKDFDPLEKIALIMSSSGTT 199  
DB 151 ----KVICVDSAP-----GCVFHSELIQSDHEITPDVKIQPDD-----VVALPYSSGTT 196  
QY 200 GLPGVVLSHRSLSLIR-----FVHSRDPYIGTRTPVQTSILSLVPPHFAFGMF 247  
DB 197 GLPGVVLTHKGLVTSVAQQVDGNGENANLYMHSD-----VLMCVLPFLPHIYSLN 245  
QY 248 TTLGVFV-VGLKVVMLKFFGALFLKTIQNYKIPTIVVAPVPMVFLAKSLVDQYDLSL 306  
DB 246 SVLLCALRVGAAILMOKFDIAQFLIPKHKVTIGTPVPIVLAIAKSPVHVYDLSV 305  
QY 307 TEVATGAPLKGKDAVAAKRLKLPG--IIQYGLTET-----CCAVMITPHNAVKTGS 358  
DB 306 RTVMSGAAPLKEEDAV--RAKFPNAKLGQGYGTEAGPVLAMCLAFAPKEPD-IKSGA 362  
QY 359 TGRPLPYIKAKVLNDATGKALGPGERGEICFQSEMIMKGYNNPEATIDTIDKGLHSG 418  
DB 363 CGTVWRNAEMKIIVDPDTCGLPRNQPGEICIRGQIMKGYLNDPEATARTIEKGLWHTG 422  
QY 419 DIGYDDEGNFFIVDRKLKELIKYGVQVAPAELENLLQHPSTADAGVTCVPDEFGQLP 478  
DB 423 DIGFIDDDDELFIYDRKLKELIKYGVQVAPAELEALLNHDPDISDAAVPMIDEAGEVP 482  
QY 479 AACVLESKTLTEKEVQDFIAAQVPTPKHLRGVVFVDSIPKGTGKLIRKELR 533  
DB 483 VAFVVRNSGSKTTEDEIKQYISKQVIFYKRIK-RVFFETVTPKSPSGKILRKDLR 536

## RESULT 15

JU0311  
4-coumarate--CoA ligase (EC 6.2.1.12) - rice  
N:Alternate names: 4-coumaroyl-CoA synthetase  
C:Species: Oryza sativa (rice)  
C:Date: 30-Jun-1992 #sequence\_revision 20-Aug-1994 #text\_change 05-May-2000  
C:Accession: JU0311  
R:Zhao, Y.; Kung, S.D.; Dube, S.K.  
Nucleic Acids Res. 18, 6144, 1990  
A:Title: Nucleotide sequence of rice 4-coumarate:CoA ligase gene, 4-CL.1.  
A:Reference number: JU0311; MUID:91045096  
A:Accession: JU0311  
A:Status: translation not shown  
A:Molecule type: DNA  
A:Residues: 1-563 <ZHA>  
A:Cross-references: EMBL:X52623; NID:g20160; PIDN:CAA36850.1; PID:g20161  
A:Experimental source: strain Japonica  
C:Genetics:  
A:Gene: 4-CL.1  
A:Introns: 351/3; 418/1; 489/2; 520/3  
C:Superfamily: 4-coumarate--CoA ligase; acetate--CoA ligase homology  
C:Keywords: acid-thiol ligase; coenzyme A; flavonoid biosynthesis  
F:87-549/Domain: acetate--CoA ligase homology <ACL>

Query Match		29.6%	Score 833.5;	DB 1;	Length 563;
Best Local Similarity		39.3%	Pred. No. 1.7e-51;		
Matches 207;		Conservative 89;	Mismatches 192;	Indels 39;	Gaps 16;
Qy	30	YKFASFPE-----AIIDAHTEVISAQIFETSCRLAVSIEQ--YGLNENNNVGVCSENN	82		
Db	42	YCFERLPEVAARPCLDIGATGVLTADVDRLSRRLAALRRAPLGRGGVVMSSLRNS	101		
Qy	83	INFPNVLAAALYLGPVATSDMTDGLTGHNLNISKPTIMFSSKKALPLILRVQOQLSF	142		
Db	102	PEFVLSFFAASRGVAATVTPMSTPHIESQLAAAGATVVITESMAAD--KLPSHSHG	158		
Qy	143	IKKVVVIDSMYDINGVECVSTFVARYTDHTFDPLSFTPKD--FDPLEKIALIMSSSGTT	199		
Db	159	ALTVVLLIDERD---GCLHFWDDLMSDEDEASPLAGDEDEKVFDPDDVVAL-PYSSGTT	213		
Qy	200	GLPKGVVLSHRSLT---IRFVHSRDPYIGTRTVPQTSILSLVPFHAFGMFTTISYFV-V	255		
Db	214	GLPKGVMLTHRSLSSTVAQQVDGENPNIGLHA--GDVILCALPMFHIYSLNTIMMCLRV	271		
Qy	256	GLKVVMLKKFSGALFLKTIQNYKIPTIVVAPPVWVFLAKSPLDVDOYDLSSLTEVATGAP	315		
Db	272	GAAIVMRRFDLAAMMDLVERHRTIAPLPPIVVAVAKSEAAAARDLSSVRMVLGAAP	331		
Qy	316	LKDVAAEAVAKRLKLPGLI--QGYGLTET-----CCAAMITPHNAVKTGSTGRPLPYIK	367		
Db	332	MKGIEDAF--WAKLPGAVLQGGYGMTEAGPVLSCLAFAKEPK-KVSGACGTVVRNAE	388		
Qy	368	AKVLDNATGKALGPGEIGEICFQSEMIMKGYNNNPEATIDTIDKDWLHSGDIGYDEDG	427		
Db	389	LKIIDPDGKSLGNLRGEICIRGQIMKGYLNNPEATKNTIDAEGLHTGDIGYVDDDD	448		
Qy	428	NFFIVDLKELIKYGYVAPAELENLLQHPSTADAGVTGVPDEFGGQLPAACVVLSESG	487		
Db	449	EIFIVDLKEIKYRGFOVAPAELEALLNTHPSIADAAVAVGL--KF-GEIPVAFVAKTEG	505		
Qy	488	KTLTEKEVODFIAAQVTPTKHLRGVVFVDSIPKGPCTGKLIRKELRE	534		
Db	506	SELSEDDVKQFVAKEVIYYKKIR-EVFFVDKIPKAPSGKIIRKELRK	551		

Search completed: September 4, 2002, 15:04:23  
Job time: 142 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 4, 2002, 15:02:46 ; Search time 39.58 Seconds  
(without alignments)  
386.478 Million cell updates/sec

Title: US-09-993-874-2  
Perfect score: 2812

Sequence: 1 MEEENIRHGERPRDIVHGPS.....KLIRKELREIFAORAPKSKL 545

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 763338 seqs, 245939087 residues

Total number of hits satisfying chosen parameters: 763338

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending\_Patents\_AA\_New.\*

- 1: /cgn2\_6/ptodata/2/paa/PCT\_NEW\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/2/paa/US06\_NEW\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/2/paa/US07\_NEW\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/2/paa/US08\_NEW\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/2/paa/US09\_NEW\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/2/paa/US10\_NEW\_COMB.pep.\*
- 7: /cgn2\_6/ptodata/2/paa/US60\_NEW\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2812	100.0	545	5	US-09-791-537-100770
2	2099.5	74.7	546	5	US-09-791-537-100774
3	1584.5	56.3	550	5	US-09-791-537-123865
4	1584.5	56.3	1172	6	US-10-122-706-4
5	1584	56.3	549	5	US-09-791-537-101683
6	1583.5	56.3	550	5	US-09-791-537-8343
7	1583.5	56.3	552	6	US-10-111-723-2
8	1568.5	55.8	550	5	US-09-791-537-124509
9	1568	55.8	551	5	US-09-791-537-124513
10	1568	55.8	552	5	US-09-791-537-8770
11	1568	55.8	553	5	US-09-791-537-38560
12	1557.5	55.4	550	5	US-09-791-537-8344
13	1534	54.6	547	5	US-09-791-537-33330
14	1519	54.0	552	5	US-09-791-537-131812
15	1516	53.9	666	5	US-09-581-894A-16
16	1516	53.9	715	5	US-09-581-894A-5
17	1515	53.9	552	5	US-09-791-537-131813
18	1515	53.9	683	5	US-09-581-894A-14
19	1515	53.9	684	5	US-09-581-894A-18
20	1515	53.9	684	5	US-09-581-894A-20
21	1511.5	53.8	548	5	US-09-791-537-27875
22	1491.5	53.0	548	5	US-09-791-537-123864
23	1485	52.8	548	5	US-09-791-537-45469
24	1483	52.7	548	5	US-09-791-537-53113
25	1458.5	51.9	548	5	US-09-791-537-73452
26	1457.5	51.8	548	5	US-09-791-537-17764

#### ALIGNMENTS

##### RESULT 1

US-09-791-537-100770  
; Sequence 100770, Application US/09791537

; GENERAL INFORMATION:

; APPLICANT: Blonmix, Inc.

; APPLICANT: Debe, Derek

; APPLICANT: Danzer, Joseph

; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME

; FILE REFERENCE: 261/210

; CURRENT APPLICATION NUMBER: US/09/791,537

; CURRENT FILING DATE: 2001-02-22

; NUMBER OF SEQ ID NOS: 153055

; SOFTWARE: Patentin version 3.0

; SEQ ID NO 100770

; LENGTH: 545

; TYPE: PRT

; ORGANISM: Phrixothrix vivianii

US-09-791-537-100770

Query Match 100.0%; Score 2812; DB 5; Length 545;  
Best Local Similarity 100.0%; Pred. No. 4.7e-235;  
Matches 545; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MEEENIRHGERPRDIVHGPSAGQQLYQSLYKFPASPEAIIDAHTNEVISYAIQIFETSCR	60
Db	1	MEEENIRHGERPRDIVHGPSAGQQLYQSLYKFPASPEAIIDAHTNEVISYAIQIFETSCR	60
Qy	61	AVSISQYGLNENNVGVCSENNINFPVLAALYLGIPVATNDMYTDGELTGHNLISKP	120
Db	61	AVSISQYGLNENNVGVCSENNINFPVLAALYLGIPVATNDMYTDGELTGHNLISKP	120
Qy	121	TFMFSKKALPLILRQOQLNLSFIKKVWVDSMYDINGVECVSTFVARYTDHTFDPLSFTF	180
Db	121	TFMFSKKALPLILRQOQLNLSFIKKVWVDSMYDINGVECVSTFVARYTDHTFDPLSFTF	180
Qy	181	KDFDPLEKIALTMSSGTTGLPKGVVLSHRSLTIRFVHSRDPYIGTRTVQTSILSLVPF	240
Db	181	KDFDPLEKIALTMSSGTTGLPKGVVLSHRSLTIRFVHSRDPYIGTRTVQTSILSLVPF	240
Qy	241	HHAFGFTTSLFVVGKLVMLKKFEGALFKTIQNKIPTIIVAPPVWVFLAKSLPVDQ	300
Db	241	HHAFGFTTSLFVVGKLVMLKKFEGALFKTIQNKIPTIIVAPPVWVFLAKSLPVDQ	300
Qy	301	YDLSSLTETVATGAPLGKDVAAEAVAKRLKLPGLIOGYGLTETCCAVMTIPTHNAVKTGSTG	360
Db	301	YDLSSLTETVATGAPLGKDVAAEAVAKRLKLPGLIOGYGLTETCCAVMTIPTHNAVKTGSTG	360

27	1454	51.7	545	5	US-09-791-537-61684	Sequence 61684, A
28	1453.5	51.7	548	5	US-09-791-537-123832	Sequence 123832, A
29	1393	49.5	543	5	US-09-791-537-60969	Sequence 60969, A
30	1388	49.4	543	5	US-09-791-537-118498	Sequence 118498, A
31	1382	49.1	543	5	US-09-791-537-60967	Sequence 60967, A
32	1377	49.0	543	5	US-09-791-537-56394	Sequence 56394, A
33	1030	36.6	544	5	US-09-791-537-8822	Sequence 8822, A
34	864.5	30.7	542	5	US-09-791-537-49392	Sequence 49392, A
35	862	30.7	547	5	US-09-791-537-19666	Sequence 19666, A
36	855	30.4	542	5	US-09-791-537-120527	Sequence 120527, A
37	847.5	30.1	636	5	US-09-791-537-84582	Sequence 84582, A
38	847	30.1	545	5	US-09-791-537-107375	Sequence 107375, A
39	847	30.1	548	5	US-09-791-537-120555	Sequence 120555, A
40	838	29.8	545	5	US-09-791-537-107378	Sequence 107378, A
41	836.5	29.7	553	5	US-09-791-537-69014	Sequence 69014, A
42	834	29.7	545	5	US-09-791-537-56966	Sequence 56966, A
43	833.5	29.6	563	5	US-09-791-537-96123	Sequence 96123, A
44	829.5	29.5	544	6	US-10-202-307-349	Sequence 349, App
45	823.5	29.3	557	5	US-09-791-537-120554	Sequence 120554, A

Qy 361 RPLPYIKAKVLDNATGKALGPCERGEICFQSEIMKGYNNPEATIDTDKDGWLHSGDI 420  
Db 361 RPLPYIKAKVLDNATGKALGPCERGEICFQSEIMKGYNNPEATIDTDKDGWLHSGDI 420  
Qy 421 GYDDEGNFFIVDRUKELIKYGVQVAPAELENLLQHPSTADAGVTGVPDEFGQLPAA 480  
Db 421 GYDDEGNFFIVDRUKELIKYGVQVAPAELENLLQHPSTADAGVTGVPDEFGQLPAA 480  
Qy 481 CVWLESGLTTEKEVODFIAAQTPTKHLRGVVFVDSIPKGTGKLRKELREIFAORA 540  
Db 481 CVWLESGLTTEKEVODFIAAQTPTKHLRGVVFVDSIPKGTGKLRKELREIFAORA 540  
Qy 541 PKSKL 545  
Db 541 PKSKL 545  
RESULT 2  
US-09-791-537-100774  
; Sequence 100774, Application US/09791537  
; GENERAL INFORMATION:  
; APPLICANT: Bionomix, Inc.  
; APPLICANT: Debe, Derek  
; APPLICANT: Danzer, Joseph  
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB  
; TITLE OF INVENTION: METHODS OF USE THEREOF  
; FILE REFERENCE: 261/210  
; CURRENT APPLICATION NUMBER: US/09/791,537  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 153055  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 100774  
; LENGTH: 546  
; TYPE: PRT  
; ORGANISM: Phrixothrix hirtus  
US-09-791-537-100774

Query Match 74.7%; Score 2099.5; DB 5; Length 546;  
Best Local Similarity 71.1%; Pred. No. 2.8e-173;  
Matches 391; Conservative 84; Mismatches 66; Indels 9; Gaps 3;  
Qy 1 MEEENIRHGERPRDIVHPGSAGQOOLYQSLYKFASPEAIIAHTNEVISAQIFETSCRL 60  
Db 1 MEEENVNGDRDLVPFGTAGLQYSLYKYSYITDGIIDAHTNEVISAQIFETSCRL 60  
Qy 61 AVSIEQYGLNENNVVGVCSENNINFFNPVLAALYLGIPVATSNMYTDCBELTGHNLNISKP 120  
Db 61 AVSLEKYGLDNHNVVAICSENNIHFFGLIAALYQGIPIMATSNMYTEREMIGHNLNISKP 120  
Qy 121 TMFSSKKALPLILRVQONLSFIKKVVVIDSMYDINGVECVSTFVARVTDHTFDPLSPTP 180  
Db 121 CLMCSKSLPLILVKQHLDFLRVVIDSMYDINGVECVFSDRSRTDHAFDVKNFP 180  
Qy 181 KDFDPLEKIALIMSSSGTGLPKGVVLSHRSLTIRFVHSRDPYIGTRVTPQTSILSLVP 240  
Db 181 KEFOPLERTALIMTSSGTTGLPKGVVLSHRSLTIRFVHSRDPYIGTRVTPQTSILAIAPF 240  
Qy 241 HHAFGMFTLSYFVVGKLVVMLKKEGALFLKTIQNYKIPTIVVAPPVWVFLAKSPVDQ 300  
Db 241 HHAFGLFTALAYFPVGLIVVWKKFEGEFLKTIQNYKIASIVVPPPIWVYLAKEPLVDE 300  
Qy 301 YDLSLSTEVATGGAPLGKDVAAEAVAKRLKPGIIOGYGLTETCCAVMITPHN-AVKTGST 359  
Db 301 YNCSSLTEIASGSGPLGRDIADKAKRLKVHGILOGYGLTETCSALILSPNDRELKKA 360  
Qy 360 GRPLPYIAKVLNATGKALGPCERGEICFQSEIMKGYNNPEATIDTDKDGWLHSGD 419  
Db 361 GTPMPYQVQKVIDINTGKALGPRESKGETCFKQMLMKGYHNNPQATDALDKDGLHTGD 420  
Qy 420 IGYDDEGNFFIVDRUKELIKYGVQVAPAELENLLQHPSTADAGVTGVPDEFGQL-- 477  
Db 421 LCYDDEGNFFIVDRUKELIKYGVQVAPAELENLLQHPNISAGV-----EFTNLLV 476

Qy 478 --PAACVVVLESGKTLTEKEVODFIAAQTPTKHLRGVVFVDSIPKGTGKLRKELREI 535  
Db 477 NYLSACVVVLEPGKTKTEKEVODYIAELVTTTKHLRGVVFIDSIPKGTGKLMRNLRAI 536  
Qy 536 FAQAPAKSKL 545  
Db 537 FAREQAKSKL 546  
RESULT 3  
US-09-791-537-123865  
; Sequence 123865, Application US/09791537  
; GENERAL INFORMATION:  
; APPLICANT: Bionomix, Inc.  
; APPLICANT: Debe, Derek  
; APPLICANT: Danzer, Joseph  
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME  
; TITLE OF INVENTION: METHODS OF USE THEREOF  
; FILE REFERENCE: 261/210  
; CURRENT APPLICATION NUMBER: US/09/791,537  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 153055  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 123865  
; LENGTH: 550  
; TYPE: PRT  
; ORGANISM: pdb 1LCI  
US-09-791-537-123865

Query Match 56.3%; Score 1584.5; DB 5; Length 550;  
Best Local Similarity 55.5%; Pred. No. 1.3e-128;  
Matches 305; Conservative 103; Mismatches 137; Indels 5; Gaps 3;  
Qy 1 MEE-ENIRHGERPRDIVHPGSAGQOOLYQSLYKFASPE--AIIAHTNEVISAQIFETS 57  
Db 1 MEDAKNIKKGPAPFPYPLEDGTAGEQLHKMKRYALVPGTIAFTDAHIEVNITYAEVFEMS 60  
Qy 58 CRLAVSIEQYGLNENNVVGVCSENNINFFNPVLAALYLGIPVATSNMYTDCBELTGHNLN 117  
Db 61 VRLAEMKRYGLNTNHRVVCSENSLSQFPMVPLGALFVGAVAPANDIYNRELLNSMNI 120  
Qy 118 SKPTIMFSSKKALPLILRVQONLSFIKKVVVIDSMYDINGVECVSTFVARVTDHTFDPLS 177  
Db 121 SQTVPVFSKGLQKLVNQQKLPKIIQIIMDSKTDYQGFQSMYTFVTSHLPPGFNEYD 180  
Qy 178 FTKDFFDPLEKIALIMSSSGTGLPKGVVLSHRSLTIRFVHSRDPYIGTRVTPQTSILSL 237  
Db 181 FVPSFDRDKTIALIMSSSGTGLPKGVVALPHRTACVRFSHARDPIFGNQIIPDTAILSV 240  
Qy 238 VPFHHAFGMFTLSYFVVGKLVVMLKKEGALFLKTIQNYKIPTIVVAPPVWVFLAKSPL 297  
Db 241 VPFHHGFMFTLGYLICGFRVLMYREBELFLRSLQDYKIQSALLVPTLFSFFAKSTL 300  
Qy 298 VDQDLSLSTEVATGGAPLGKDVAAEAVAKRLKPGIIOGYGLTETCCAVMITPHNAVKTG 357  
Db 301 IDKYDLSNLHEIASGAPLSKEVGEAVAKRFLPGIRGYGLTETTSAILITPEGDDRP 360  
Qy 358 STGRPLPYIAKVLNATGKALGPCERGEICFQSEIMKGYNNPEATIDTDKDGWLHLS 417  
Db 361 AVGVVVPFEAKVVDLDTGKTLGVNQRGELCVRGPMIMSGVYNNPEATNALIDKDWLHS 420  
Qy 418 GDIGYDDEGNFFIVDRUKELIKYGVQVAPAELENLLQHPSTADAGVTGVPDEFGQL 477  
Db 421 GDIAWDEDEHFFIVDRUKSLIKYGVQVAPAELESILLQHPNIPFDAGVAGLPDDADAGEL 480  
Qy 478 PAACVVVLESGKTLTEKEVODFIAAQTPTKHLRGVVFVDSIPKGTGKLRKELREIF- 536  
Db 481 PAAVVVLEHGKTMTEKEIVDVVASQVTTAKKLRGVVFVDEVPKGLTCKLDARKIREILI 540  
Qy 537 -AQRAPAKSKL 545  
Db 537 -AQRAPAKSKL 545





; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 153055  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 8343  
; LENGTH: 550  
; TYPE: PRT  
; ORGANISM: Photinus pyralis  
US-09-791-537-8343

Query Match 56.3%; Score 1583.5; DB 5; Length 550;  
Best Local Similarity 55.5%; Pred. No. 1.6e-128;  
Matches 305; Conservative 102; Mismatches 138; Indels 5; Gaps 3;

Qy 1 MEE-ENIRHGERPRDIVHPSAGQOLYQSLYKFAFPF--AIIAHTNEVISYAQIFETS 57  
Db 1 MEDAKNIKKGPAPFYPLEDGTAGEQLHKAMKRYALVPGTIAFTDAHIEVNITYAEYFEMS 60  
Qy 58 CRLAVSIEQYGLNENNVVGVCSENNINFFNPVLAALYLGIPVATSNOMYTDCGELTGHINI 117  
Db 61 VRLAEAMKRYGLNTHNRIVWCSENSLOPFMPVLGALFIGVAVAPANDIYNRELLNSMNI 120  
Qy 118 SKPTIMSSKALPLILRVQONLSFIKKVVVIDSMYDINGVECVSTFVARTDHTFDPLS 177  
Db 121 SQTVPVFSKKGLOKILNVOKKLPPIQKIIIMDSKTDYQGFQSMYTFVTSHLPFCFNEYD 180  
Qy 178 FTKDFDPLEKIALIMSSGTTGLPKGVLSHRSLTIRFVHSRDIPIYGRTPVQTSILSL 237  
Db 181 FVPESFDRDKTIALIMSSGTTGLPKGVLPHTACVRFSHARDPIFGNQIIPDTAILSV 240  
Qy 238 VPFHAFGMFTLSYFVVGVLKVMKKEGALFLKTIQNYKIPTIVVAPVVMVFLAKSPL 297  
Db 241 VPFHFGFMFTLGLYICGFRVWLMYRFEELFLRSLODYKIQSALLVPTLFSFFAKSTL 300  
Qy 298 VDQYDLSLSEVATGGAPLGKDVAEAVAKRLKPGIIOGYGLTETCCAVMITPHNAVKTG 357  
Db 301 IDKYDLSNLHEIASGGAPLSKEVGEAVAKRPHLPICRQGYGLTETTSAILITPKGDFKPG 360  
Qy 358 STGRPLPIYAKVLDNATGKALGPGEICEICFQSEMIMKGYNNPEATIDIDKDGWLHS 417  
Db 361 GVGKVPVFEAKVVDLTGKTLGVNQRGELCVRGPMINSGYVNNPEATNALIDKDGWLHS 420  
Qy 418 GDIGYDEGDNFFIVDRUKELIKYGVQVAPAELENLLQHPSTADAGVTCVDFEGQOL 477  
Db 421 GDIAVWDEDEHFFIVDRUKLSIKYGVQVAPAELESILLQHPNIFDAGVAGLPDDDAGEL 480  
Qy 478 PAACVWLESKGLTEKEVQDFIAAQVPTKHLRGVGVFVDSIPKGPCKLIRKELREIF- 536  
Db 481 PAAVVLHGKMTKEIEVDYVASQVTTAKKLRGGVVFVDEVPKGLTGKLDARKIREILI 540  
Qy 537 -AORAPSKL 545  
Db 541 KAKKGKSKL 550

## RESULT 7

US-10-111-723-2  
; Sequence 2, Application US/10111723  
; GENERAL INFORMATION:  
; APPLICANT: The Secretary of State for Defence  
; APPLICANT: White, Peter J  
; APPLICANT: Willey, Tara L  
; APPLICANT: Price, Rachel L  
; APPLICANT: Murphy, Melanie J  
; APPLICANT: Squirell, David  
; TITLE OF INVENTION: Novel Enzyme  
; FILE REFERENCE: DERA/IPD/PI247/WOD  
; CURRENT APPLICATION NUMBER: US/10/111.723  
; PRIOR FILING DATE: 2002-04-26  
; PRIOR APPLICATION NUMBER: GB 9925161.3  
; PRIOR FILING DATE: 1999-10-26  
; PRIOR APPLICATION NUMBER: GB 0016744.5  
; PRIOR FILING DATE: 2000-07-10

; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 552  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Encoded amino acid  
; OTHER INFORMATION: sequence of the cDNA sequence of synthetic  
; OTHER INFORMATION: luciferase gene  
US-10-111-723-2

Query Match 56.3%; Score 1583.5; DB 6; Length 552;  
Best Local Similarity 55.5%; Pred. No. 1.6e-128;  
Matches 305; Conservative 102; Mismatches 138; Indels 5; Gaps 3;

Qy 1 MEE-ENIRHGERPRDIVHPSAGQOLYQSLYKFAFPF--AIIAHTNEVISYAQIFETS 57  
Db 3 MEDAKNIKKGPAPFYPLEDGTAGEQLHKAMKRYALVPGTIAFTDAHIEVNITYAEYFEMS 62  
Qy 58 CRLAVSIEQYGLNENNVVGVCSENNINFFNPVLAALYLGIPVATSNOMYTDCGELTGHINI 117  
Db 63 VRLAEAMKRYGLNTHNRIVWCSENSLOPFMPVLGALFIGVAVAPANDIYNRELLNSMNI 122  
Qy 118 SKPTIMSSKALPLILRVQONLSFIKKVVVIDSMYDINGVECVSTFVARTDHTFDPLS 177  
Db 123 SQTVPVFSKKGLOKILNVOKKLPPIQKIIIMDSKTDYQGFQSMYTFVTSHLPFCFNEYD 182  
Qy 178 FTKDFDPLEKIALIMSSGTTGLPKGVLSHRSLTIRFVHSRDIPIYGRTPVQTSILSL 237  
Db 183 FVPESFDRDKTIALIMSSGTTGLPKGVLPHTACVRFSHARDPIFGNQIIPDTAILSV 242  
Qy 238 VPFHAFGMFTLSYFVVGVLKVMKKEGALFLKTIQNYKIPTIVVAPVVMVFLAKSPL 297  
Db 243 VPFHFGFMFTLGLYICGFRVWLMYRFEELFLRSLODYKIQSALLVPTLFSFFAKSTL 302  
Qy 298 VDQYDLSLSEVATGGAPLGKDVAEAVAKRLKPGIIOGYGLTETCCAVMITPHNAVKTG 357  
Db 303 IDKYDLSNLHEIASGGAPLSKEVGEAVAKRPHLPICRQGYGLTETTSAILITPKGDFKPG 362  
Qy 358 STGRPLPIYAKVLDNATGKALGPGEICEICFQSEMIMKGYNNPEATIDIDKDGWLHS 417  
Db 363 AVGVVPEFEAKVVDLTGKTLGVNQRGELCVRGPMINSGYVNNPEATNALIDKDGWLHS 422  
Qy 418 GDIGYDEGDNFFIVDRUKELIKYGVQVAPAELENLLQHPSTADAGVTCVDFEGQOL 477  
Db 423 GDIAVWDEDEHFFIVDRUKLSIKYGVQVAPAELESILLQHPNIFDAGVAGLPDDDAGEL 482  
Qy 478 PAACVWLESKGLTEKEVQDFIAAQVPTKHLRGVGVFVDSIPKGPCKLIRKELREIF- 536  
Db 483 PAAVVLHGKMTKEIEVDYVASQVTTAKKLRGGVVFVDEVPKGLTGKLDARKIREILI 542  
Qy 537 -AORAPSKL 545  
Db 543 KAKKGKSKL 552

## RESULT 8

US-09-791-537-124509  
; Sequence 124509, Application US/09791537  
; GENERAL INFORMATION:  
; APPLICANT: Bionomix, Inc.  
; APPLICANT: Debe, Derek  
; APPLICANT: Danzer, Joseph  
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME  
; FILE REFERENCE: 261/210  
; CURRENT APPLICATION NUMBER: US/09/791.537  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 153055  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 124509





Db 241 VPFHGFHFTTLGVLICGFRVVLMYRFEELFLRLSDQYKIQSALLVPTLFSFAKSTL 300  
QY 298 VDOYDLSSLETEVATGAGPLGKDAEAVAKRLKLPGLIGQYGLTETCCAVMITHPNHAKVTG 357  
Db 301 IDKYDLSNLHEIASGAPLSKEVGEAVAKRFLPLGIRQYGLTETTSAILITPEGDDKPG 360  
QY 358 STGRPLPYTKAKVLDNATKALGPGERIGCFQSEMIMKGYNNPEATIDTIDKDGWLHS 417  
Db 361 AVGVVPPFEAKVVDLDTGTLGNQVORGLCVGRPMIMSGYNDPEATNALIDKDGWLHS 420  
QY 418 GOIGYDEGNFFIVDRKLKELIKYGYQVAPAELENLLQHPHSIADAGVTGVPDEFEGGOL 477  
Db 421 GDIATWDEHEFFIVDRKLSLKIYGCQVAPAELESILLOHPHFIDAGVAGLPDODAGEL 480  
QY 478 PAACVVLESKTLTEKEVODFIAAQTPTKHLRGVGVFVDSIPKPGTGLIRKELREIF- 536  
Db 481 PAAVVVLEHGKMTKEIVDYVASQVTTAKLRLGGVGVFVDEVPKGLTGKLDARKIREIL 540  
QY 537 -AORAPKSKL 545  
Db 541 KAKKGKSKL 550

RESULT 13  
US-09-791-537-33330  
; Sequence 33330, Application US/09791537  
; GENERAL INFORMATION:  
; APPLICANT: Blonomix, Inc.  
; APPLICANT: Debe, Derek  
; APPLICANT: Danzer, Joseph  
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS  
; FILE REFERENCE: 261/210  
; CURRENT APPLICATION NUMBER: US/09/791,537  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 153055  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 33330  
; LENGTH: 547  
; TYPE: PRT  
; ORGANISM: Lampyris noctiluca  
US-09-791-537-33330

Query Match 54.6%; Score 1534; DB 5; Length 547;  
Best Local Similarity 53.5%; Pred. No. 3.1e-124;  
Matches 293; Conservative 106; Mismatches 145; Indels 4; Gaps 3;

QY 1 MEE-ENIRHGERPRDIVHPGSGAQOQLYQSLYKFAFPE--AIIDAHTNEVSYAQIFETS 57  
Db 1 MEDAKNIMHGPAFFYPLEBDGTAGEQLHAKMKRYAQPVGTAFTDAHAENVITYSEYFEMA 60  
QY 58 CRLAVSIEQYGLNENNVGVCSENNINFPNVLAAALYLGIPVATSNDMYTDGELTCHLNI 117  
Db 61 CRLAETMKRYGLQHLHIAVCSENSLOFPNVCAGLFIGVGVASTNDIYNRELNSLSI 120  
QY 118 SKPTIMFSSKALPLILRVQOQLSFYKVVVWIDSMYDINGVECVSTFVARYTDHTFDPLS 177  
Db 121 SQTIVSCSKRALQKILGVOKKLPILQIKVILDSREDYMGKQSMYSFIESHLPAAGNEYD 180  
QY 178 FYPKDFDPLEKTALIMSSSGTGLPKGVVLSHRSLTIRVHSDRPIYGRTRTPQTSILS 237  
Db 181 YIPDSFRETATALIMSSSGTGLPKGVVLSHRSLTIRVHSDRPIYGRTRTPQTSILS 240  
QY 238 VPFHAFGFTTSLYFVGLKVVMLKKFEGALFLKTIQNKIPTIVAPPVMMFLAKSP 297  
Db 241 IFPHGFGFTTLGYLTCGFRVVLMYRFEELFLRLSDQYKIQSALLVPTLFSFAKSTL 300  
QY 298 VDOYDLSSLETEVATGAGPLGKDAEAVAKRLKLPGLIGQYGLTETCCAVMITHPNHAKVTG 357  
Db 301 VDKYDLSNLHEIASGAPLAKEVGEAVAKRFLPLGIRQYGLTETTSAILITPEGDDKPG 360  
QY 358 STGRPLPYTKAKVLDNATKALGPGERIGCFQSEMIMKGYNNPEATIDTIDKDGWLHS 417

Db 361 ACCKVVPFESAKIVDLDTGKTLGVNORGLCVKGMIMKGYNNPEATNALIDKDGWLHS 420  
QY 418 GOIGYDEGNFFIVDRKLKELIKYGYQVAPAELENLLQHPHSIADAGVTGVPDEFEGGOL 477  
Db 421 GDIATWDEHEFFIVDRKLSLKIYGCQVAPAELESILLOHPHFIDAGVAGLPDODAGEL 480  
QY 478 PAACVVLESKTLTEKEVODFIAAQTPTKHLRGVGVFVDSIPKPGTGLIRKELREIFA 537  
Db 481 PAAVVVLEHGKMTKEIVDYVASQVTTAKLRLGGVGVFVDEVPKGLTGKLDARKIREIL- 539  
QY 538 QORAPKSKL 545  
Db 540 MMGKSKL 547

RESULT 14  
US-09-791-537-131812  
; Sequence 131812, Application US/09791537  
; GENERAL INFORMATION:  
; APPLICANT: Blonomix, Inc.  
; APPLICANT: Debe, Derek  
; APPLICANT: Danzer, Joseph  
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS  
; FILE REFERENCE: 261/210  
; CURRENT APPLICATION NUMBER: US/09/791,537  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 153055  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 131812  
; LENGTH: 552  
; TYPE: PRT  
; ORGANISM: Photuris pennsylvanica  
US-09-791-537-131812

Query Match 54.0%; Score 1519; DB 5; Length 552;  
Best Local Similarity 52.8%; Pred. No. 6.2e-123;  
Matches 287; Conservative 102; Mismatches 151; Indels 4; Gaps 3;

QY 3 ENIRHGERPRDIVHPGSGAQOQLYQSLYKFAFPE--AIIDAHTNEVSYAQIFETS 60  
Db 4 ENNLTGPPPPYPLEEGTAGEQHLRAITRYAAVPGTAYTDVHTLEVITYKEFLDVTCL 63  
QY 61 AVSIEQYGLNENNVGVCSENNINFPNVLAAALYLGIPVATSNDMYTDGELTCHLNI 120  
Db 64 AAMKNGYGLQHTISVCSENCVQFPMPCAAALYIGVATAPNDIYNRELNSLSISQP 123  
QY 121 TMFSSKALPLILRVQOQLSFYKVVVWIDSMYDINGVECVSTFVARYTDHTFDPLSFT 180  
Db 124 TVVFTSRNSLQKILGVQSLPVIKKIIMLDTKDYLGYSQSMOSFMKEHPANFNVSAPK 183  
QY 181 KDFDPLEKTALIMSSSGTGLPKGVVLSHRSLTIRVHSDRPIYGRTRTPQTSILS 240  
Db 184 LSFD-LDRVACIMNSSSGTGLPKGVDPISHRNTYRSHCRDPVFGNQIIPDTTILCAV 242  
QY 241 HIAFGFTTSLYFVGLKVVMLKKFEGALFLKTIQNKIPTIVAPPVMMFLAKSP 300  
Db 243 HIAFGFTTSLYFVGLKVVMLKKFEGALFLKTIQNKIPTIVAPPVMMFLAKSP 302  
QY 301 YDLSSLETEVATGAGPLGKDAEAVAKRLKLPGLIGQYGLTETCCAVMITHPNHAKVTG 360  
Db 303 YDLSSLETEVATGAGPLGKDAEAVAKRLKLPGLIGQYGLTETCCAVMITHPNHAKVTG 362  
QY 361 RELPYTKAKVLDNATKALGPGERIGCFQSEMIMKGYNNPEATIDTIDKDGWLHS 420  
Db 363 KVVFPYSLKVLDTNLTKGLGNPNERGICFTGMIMKGYNNPEATREIDEEGWIHS 422  
QY 421 GYDEGNFFIVDRKLKELIKYGYQVAPAELENLLQHPHSIADAGVTGVPDEFEGGOL 480  
Db 423 GYDEGNFFIVDRKLKELIKYGYQVAPAELENLLQHPHFIDAGVAGVPDEFAGDLP 482

Qy 481 CVMLESGKLTKEKEVQDFIAAQVTTKHLRGVGVFVDSIPKPTGKTLRKELREIFAORA 540  
Db 483 VVVLKRGSKSITEKEIQDYVAGOVTSKKLRGGVEFVKPVGFTGKIDTRKIKEILI-KA 541  
Qy 541 PKSK 544  
Db 542 QXGK 545  
RESULT 15  
US-09-581-894A-16  
; Sequence 16, Application US/09581894A  
; GENERAL INFORMATION:  
; APPLICANT: Fong, Tung Ming  
; APPLICANT: van der Ploeg, Leonardus H. T.  
; APPLICANT: Tota, Michael R.  
; TITLE OF INVENTION: C-TERMINAL REGION OF AGOUTI-RELATED  
; TITLE OF INVENTION: TRANSCRIPT (ART) PROTEIN  
; FILE REFERENCE: 20146P  
; CURRENT APPLICATION NUMBER: US/09/581.894A  
; CURRENT FILING DATE: 2000-08-21  
; PRIOR APPLICATION NUMBER: PCT/US98/26457  
; PRIOR FILING DATE: 1998-12-11  
; PRIOR APPLICATION NUMBER: 60/069,747  
; PRIOR FILING DATE: 1997-12-16  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 16  
; LENGTH: 666  
; TYPE: PRT  
; ORGANISM: Human  
US-09-581-894A-16

Query Match 53.9%; Score 1516; DB 5; Length 666;

Best Local Similarity 51.4%; Pred. No. 1.5e-122;  
Matches 286; Conservative 105; Mismatches 151; Indels 14; Gaps 3;

Qy 3 EENIRHGERPRDIVHPGAGQOLYQSLYKFAFPE--AIIDAHTNEVISYAOIFETSCRL 60  
Db 95 ENNILLGPPPPYPLEGEGTAGELHRAISRYAAVPGTLAYTDVHTLEVTYKEFLDVTCL 154  
Qy 61 AVSIEOYGLNENNVGVCSENNINFPNVLAAALYLGIPVATSNMDYTDGELTGHNLISKP 120  
Db 155 AEAMKNYGLGQHTISVCSENCVPFMPICAAALYGVATAPTNDIYNRELYNSLSISQP 214  
Qy 121 TTFSSKALPLILRVOONLSPIKKVWIDSNYDINGECVSTFVARYTDHTFDPLSETP 180  
Db 215 TVVFTSRNSLQKILGVQSRPLPIKKIIILDGKKDYLGYQSMQSEFMKEHVPANFNVSAPKP 274  
Qy 181 KDFDPLEKIALIMSSGGTGLPKGVVLSHRSLTIRFVHSRDPYIGTRTPQTSILSLVPF 240  
Db 275 LSFQ-LDRVACIMNMSGTGLPKGVPISHRNTIYRFSHCRDPVFGNQIIPDTTILCAVPF 333  
Qy 241 HFAFGFTTSLYFVGLKVVMLKKEGALFLKTIQNYKIPTIVVAPPVWFLAKSPLVDQ 300  
Db 334 HFAFGFTNLGLIGGFHVLMYRENEHLFLQTDYKCSALLVPTVLAFIAKNPLVDK 393  
Qy 301 YDLSSITEVATCGAPLGRDVAEAVAKRLKLPGLIOGYGLTETCCAVMITPHNAVKTGTG 360  
Db 394 YDLSNLHEIASGGAPLSKEISEIAAKRFLPGIRGQYGLTETTCATVITAEGEFKLGAVG 453  
Qy 361 RPLPYIKARVLNATGKALPGERGEICFQSEMIMKGYNNPEATIDTIDKDWLHSGDI 420  
Db 454 KVPFYSKLVLDLNTGKGLGPNERGEICFKGPMIMKGYNNPEATRELIDEGWIHSGDI 513  
Qy 421 GYDEGDNFFIVDRKELIKYGYQVAPAELENLLQHPISADAGVTGVPDEFGQLPAA 480  
Db 514 GYFDEGHHYIVVDRKLSLIKYGYQVPAELEALLQHPFIEDAGVAGVDEVAGDLPCA 573  
Qy 481 CVMLESGKLTKEKEVQDFIAAQVTTKHLRGVGVFVDSIPKPTGKTLRKELREIFA--- 536  
Db 574 VVVLKRGSKSITEKEIQDYVAGOVTSKKLRGGVEFVKPVGFTGKIDTRKIKEILIRAQ 633

Qy 537 -----AORAPKSKL 545  
Db 634 KGKSKAKLGPQK 649

Search completed: September 4, 2002, 15:05:10  
Job time: 144 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 4, 2002, 15:00:26 ; Search time 34.33 Seconds  
(without alignments)  
1763.333 Million cell updates/sec

Title: US-09-993-874-2  
Perfect score: 2812  
Sequence: 1 MEEENRHGCRPDIVHPCS.....KLIRKELREIFAQAPKSKL 545

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_032802.\*

```
1: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1588.5	56.5	550	16	AA866627
2	1585.5	56.4	550	17	AA866627
3	1584.5	56.3	550	15	AA850011
4	1584.5	56.3	550	20	AA850011
5	1584.5	56.3	550	22	AA831172
6	1583.5	56.3	550	15	AA83006
7	1583.5	56.3	550	17	AA83006
8	1583.5	56.3	552	22	AA835498
9	1581.5	56.2	815	13	AA828127
10	1577.5	56.1	743	15	AA854867
11	1577.5	56.1	743	16	AA872801

ID	Score	Query Match	Length	ID	Description
12	1576.5	56.1	550	20	AA866627
13	1568.5	55.8	550	18	AA866627
14	1568.5	55.8	550	22	AA866627
15	1568	55.8	551	18	AA866627
16	1524.5	54.2	552	20	AA866627
17	1521.5	54.1	552	20	AA866627
18	1516	53.9	666	20	AA866627
19	1516	53.9	715	20	AA866627
20	1515	53.9	552	16	AA866627
21	1515	53.9	683	20	AA866627
22	1515	53.9	684	20	AA866627
23	1515	53.9	684	20	AA866627
24	1493.5	53.1	548	18	AA866627
25	1492.5	53.1	548	18	AA866627
26	1492.5	53.1	552	20	AA866627
27	1491.5	53.0	548	10	AA866627
28	1491.5	53.0	548	11	AA866627
29	1491.5	53.0	548	22	AA866627
30	1490.5	53.0	548	18	AA866627
31	1488.5	52.9	548	18	AA866627
32	1487.5	52.9	548	14	AA866627
33	1487.5	52.9	548	14	AA866627
34	1486.5	52.9	548	12	AA866627
35	1485.5	52.8	548	12	AA866627
36	1485.5	52.8	548	12	AA866627
37	1485	52.8	548	16	AA866627
38	1484.5	52.8	548	12	AA866627
39	1483.5	52.8	548	12	AA866627
40	1483.5	52.8	548	12	AA866627
41	1482	52.7	544	22	AA866627
42	1481.5	52.7	548	12	AA866627
43	1477	52.5	544	22	AA866627
44	1475	52.5	544	22	AA866627
45	1475	52.5	546	20	AA866627

#### ALIGNMENTS

RESULT 1

AA866627

ID AA866627 standard; Protein; 550 AA.

AC AA866627;

DT 12-FEB-1996 (first entry)

XX Firefly luciferase (A215L,E354X) engineered mutant.

XX Luciferase; thermostability; enzyme stabilization; firefly;

KW Photinus pyralis; luc gene; site-directed mutagenesis; label;

KW ATP analysis.

XX Synthetic.

XX Key

PH Key

FT Misc-difference 354

FT /note= "amino acid at position 354 is any amino acid

FT other than Glu, pref. also not Gly, Pro

FT or Asp, and is pref. Trp, Val, Leu, Ile,

FT Asn or esp. Arg or Lys, their analogues

FT or modified forms"

XX WO9525798-A1.

XX 28-SEP-1995.

XX 22-MAR-1995; 95WO-GB00629.

XX 20-JAN-1995; 95GB-0001170.

XX 23-MAR-1994; 94GB-0005750.

XX (MINA ) UK SEC FOR DEFENCE.

PA

```
XX PI Lowe CR, Murray JAH, Squirrel DJ, White PJ;
XX DR WPT; 1995-344619/44.
XX PT New heat stable luciferase mutants - related DNA, vectors and
XX PT transformed cells, useful as reporter genes, assay labels and for
XX PT determ of ATP
XX PS Disclosure; Page 23-25; 40pp; English.
XX CC A doubly mutated luciferase of Photinus pyralis has leucine at
XX CC amino acid position 215 and an amino acid other than glutamic
XX CC acid at position 354. The luciferase, obtd. by site-directed
XX CC mutagenesis of the luc gene, shows improved thermostability
XX CC compared to the wild-type.
XX SQ Sequence 550 AA;

Query Match 56.4%; Score 1588.5; DB 16; Length 550;
Best Local Similarity 55.6%; Pred. No. 7e-144;
Matches 306; Conservative 103; Mismatches 136; Indels 5; Gaps 3;

QY 1 MEE-ENIRHGERPRDIVHPGSAQOOLYQSLYKFAFPE--AIIDAHTNEVISYAQIFETS 57
Db 1 medaknikkgpapfypdgedtagedqlhkamkryalvpgtiaftdahievntyaeyfems 60
QY 58 CRLAVSIQOYGLNENNVGVCSENNINFPNVLALYLGIPVATSNMYTDCGLTGHNLNI 117
Db 61 vrleamkryginthrivvcsenslqfmpvlgalfgavavapandiynerellnsnni 120
QY 118 SKPTIMFSKKALPILRLVQOQLSPFKKVVWIDSMWDINGVECVSTFVARVTDHTFDPLS 177
Db 121 sqptvvfsvkqglqklnvqkklpliqkiiimdstdyqgfgsmvtfvtshlppgfned 180
QY 178 FTPKDFDPLEKIALIMSSSGTGLPKGVVLSHRSUTIRFVHSRDPYIGTRVPTQTSILSL 237
Db 181 fvpesfdrdktlialimnssgstgipkvalphrticvrfshardpifgnqilpdtailav 240
QY 238 VPFHAFGMFTTLSYFVVGLKVMMLKKEGALFLKTIONYKIPTIVAVPPVWVFLAKSPL 297
Db 241 vpfhgfmgfttlylilcgfrvmyrfexelfrlsldqyqksallvptlfsfakatl 300
QY 298 VDQYDLSLITEVATGAPLGDVAAVAKRLKLPGLIOGYGLTETCCAVMITPHNAVKTG 357
Db 301 idkydlsnlheiasgagplskeveavakrfhlpgirgyltetsailtpxgddkpg 360
QY 358 STGRPLPYIKAKVLDNATGKALGPGERGEICFQSEMIKGYNNPEATIDTIDKDGWLHS 417
Db 361 avgvkvpfreakvvdldtktlgvqnrgelcvrgpmimgvynpneatnalldkdgwlhs 420
QY 418 GDIGYVDEGNFFIVDRLEKELIKYGYOVAEAELENLLOHPSIADAGVTGVPDEFEGOL 477
Db 421 gdiaywdehffvdrilkslikygyvapaetesilqbnifdgaagvlpdddagel 480
QY 478 PAACVLESKGLTLEKEYQDFIAQVTPTKHLRGGWFWVDSIPKPTGKLRKELREIF- 536
Db 481 peavvvlehgkmtkeivdyvaegvttakkirgvgvfvdevpkaltkldarkirelli 540
QY 537 -AQRAPKSKL 545
Db 541 kakkggkaki 550

RESULT 2
AAR98515
ID AAR98515 standard; Protein; 550 AA.
XX AC AAR98515;
XX DT 11-MAR-1997 (first entry)
XX
```

```
DE P. pyralis variant luciferase E270X.
XX XX Luciferase; firefly; ATP; heat stability; luciferin; light.
XX OS Photinus pyralis.
XX PH Key Location/Qualifiers
XX FT Misc-difference 270
XX FT /label= E270X
XX FT /note= "Mutated to improve heat stability"
XX PN WO9622376-A1.
XX PD 25-JUL-1996.
XX PF 19-JAN-1996; 96WO-GB000099.
XX PR 24-APR-1995; 95GB-0008301.
XX PR 20-JAN-1995; 95GB-0001172.
XX PA (MINA ) UK SEC FOR DEFENCE.
XX PI Lowe CR, Murray JA, Squirrel DJ, White PJ;
XX DR WPT; 1996-354533/35.
XX DR N-PSDB; NAT36492.
XX PT Mutant luciferase(s) with greater heat stability than wild-types -
XX PT useful in luminescent reagents
XX PS Claim 9; Page 20-24; 39pp; English.
XX CC This sequence represents a variant luciferase from the firefly, P.
XX CC pyralis. In this sequence the amino acid at position 270 has
XX CC been mutated such that it is not Glu. The luciferase has a
XX CC lower Km for the substrate ATP than native luciferase. The
XX CC luciferase therefore has a greater heat stability than wild-type
XX CC luciferase. The luciferase may be used in an assay to measure ATP
XX CC with luciferin, whereby the amount of light is related to the amount
XX CC of ATP present.
XX SQ Sequence 550 AA;

Query Match 56.4%; Score 1585.5; DB 17; Length 550;
Best Local Similarity 55.5%; Pred. No. 1.4e-143;
Matches 305; Conservative 103; Mismatches 137; Indels 5; Gaps 3;

QY 1 MEE-ENIRHGERPRDIVHPGSAQOOLYQSLYKFAFPE--AIIDAHTNEVISYAQIFETS 57
Db 1 medaknikkgpapfypdgedtagedqlhkamkryalvpgtiaftdahievntyaeyfems 60
QY 58 CRLAVSIQOYGLNENNVGVCSENNINFPNVLALYLGIPVATSNMYTDCGLTGHNLNI 117
Db 61 vrleamkryginthrivvcsenslqfmpvlgalfgavavapandiynerellnsnni 120
QY 118 SKPTIMFSKKALPILRLVQOQLSPFKKVVWIDSMWDINGVECVSTFVARVTDHTFDPLS 177
Db 121 sqptvvfsvkqglqklnvqkklpliqkiiimdstdyqgfgsmvtfvtshlppgfned 180
QY 178 FTPKDFDPLEKIALIMSSSGTGLPKGVVLSHRSUTIRFVHSRDPYIGTRVPTQTSILSL 237
Db 181 fvpesfdrdktlialimnssgstgipkvalphrtacvrfshardpifgnqilpdtailav 240
QY 238 VPFHAFGMFTTLSYFVVGLKVMMLKKEGALFLKTIONYKIPTIVAVPPVWVFLAKSPL 297
Db 241 vpfhgfmgfttlylilcgfrvmyrfexelfrlsldqyqksallvptlfsfakatl 300
QY 298 VDQYDLSLITEVATGAPLGDVAAVAKRLKLPGLIOGYGLTETCCAVMITPHNAVKTG 357
Db 301 idkydlsnlheiasgagplskeveavakrfhlpgirgyltetsailtpgddkpg 360
QY 358 STGRPLPYIKAKVLDNATGKALGPGERGEICFQSEMIKGYNNPEATIDTIDKDGWLHS 417
```



Db	361	avgkvvpff	feakvvdtdgtklgnvngelcvrpgm	sgvnnpeatn	ldkdgwlhs	420									
Qy	418	GDIGYYDEDCNFFIVDR	KLRIKYGQVAPAELENLLQHP	SIADAGVTGVPDEF	GGQL	477									
Db	421	gdlaywde	ehffivdrklklygvapaelesillqb	pnifdagv	aglpdddagel	480									
Qy	478	PAACVVLES	GKTLTEKEVQDFIAAQVTP	PTKHLRGVVVD	SIKPGPTGKLI	RELREIF- 536									
Db	481	paavvle	hgkmtkeivdyvasqv	ctakkl:ggvvf	vdvpgkl	egkidarkirell 540									
Qy	537	-AORAPKSL	545												
Db	541	kakkggkskl	550												
RESULT 3															
XX	ID	AAR50011	standard; Protein; 550 AA.												
XX	AC	AAR50011;													
XX	DT	29-SEP-1994	(first entry)												
XX	XX	Firefly luciferase reporter protein.													
DE	XX	Multicistronic expression unit; bicistronic vector system;													
XX	XX	recombinant protein production; reporter protein; luciferase;													
KW	XX	firefly.													
XX	XX	Photinus pyralis.													
OS	XX	W09405785-A.													
PN	XX	17-MAR-1994.													
PD	XX	26-AUG-1993; 93WO-EP02294.													
PP	XX	27-AUG-1992; 92DE-4228458.													
PR	XX	(BETE ) BEIERSDORF AG.													
XX	XX	(GBFB ) GBF GES BIOTECH FORSCHUNG GMBH.													
PA	XX	Achterberg V, Dirks W, Dorschner A, Eichner W, Hauser H;													
PI	XX	Meyer-Ingold W, Mielke H, Wirth M, Doerschner A;													
PI	XX	WPI; 1994-101190/12.													
DR	XX	N-PSDB; AAQ58732.													
DR	XX	New multicistronic expression units - for producing equimolar													
PPT	XX	amts. of polypeptide(s) in mammalian cells as hosts													
PPT	XX	Example 1; Page 66-67; 109pp; German.													
XX	XX	Reporter genes coding for secretory alkaline phosphatase (AAQ58731)													
XX	XX	and for luciferase (AAQ58732) can be co-expressed using a bicistronic													
CC	XX	system in which a sequence responsible for internal translation													
CC	XX	start (designated "IRES") is located between cistrons coding for													
CC	XX	the two proteins.													
XX	XX	Sequence 550 AA;													
XX	XX														
Query Match 56.3%; Score 1584.5; DB 15; Length 550;															
Best Local Similarity 55.5%; Pred. No. 1.7e-143;															
Matches 305; Conservative 103; Mismatches 137; Indels 5; Gaps 3;															
QY	1	MEE-ENIRGERPRDI	VHPSAGQOLYKSFASFP	E--AIDAHTNEV	ISYAQIFETS	57									
Db	1	medaknikk	gapfpledgtageqlbkamkryalvpgt	iaftdahtev	nityae	yfems 60									
QY	58	CRLAIVSEIQYGLN	ENVVGYCSENNINFFNPVLAALY	LGIPVAT	SNDMYT	DGELTGLHNI 117									

Db	61	vrlaeamkrygIntnhriivcsenslqffempvlgalfigvavapandlynereIlnsmi	120
Qy	118	SKETIMFSKKKALLIRVQOONLSFIKKVVIDSMYDINGVECVSTFVARVYTDHTFDPLS	177
Db	121	sqptvfvsvskgqlqkIlilnqkklplliqkiIlnsdsktdyggfqgmytftshlppgfneyd	180
Qy	178	FTPKDFDPLEKIALIMSSGTTGLPKGVVLHSRLTIRPVIHSRDPDIYTRIVPOTSILSL	237
Db	181	fvpesfdrdktialImssgstglpkvvalphrtacvrfshardpifgnqilpdtailsv	240
Qy	238	VPHHAFAGFTTISYFVVGILKVMILKKPKGALFLKTIQNYKIPTIVVAPVMVFLAKSPL	297
Db	241	vpfhgfmfttlgyllicfrvvlmyrfeeeelfrlsgdykIqsalIvptlfsfakatl	300
Qy	298	VDQYDLSLSTEATGGAPlGKDVAAEAVAKRLKLPGLIIQGYGLTETCCAVMITPHNAVKTG	357
Db	301	IdkydIlnhelasggapIskevegeavakrflhplgIrqgygitettsailitpegddkpg	360
Qy	358	STGRPLPYIKAKVLDNATKALGPGERGICFQSEMIKKGYNNPEATIDIDKDWLHS	417
Db	361	avgkvvpvffeakvvdIdtktIgvnqrgeIcIvrgpmImsgyvvnnpeatIIdkdgwlhs	420
Qy	418	GDIGYDDEGNFFIVDRKLKELIKYGVQVAPAELENLLOHPSIADAGVTGVPDEFGGOL	477
Db	421	gdIaydwdehffivdrIksllkyqyvapaalesillqhnifdaevagIpdddagel	480
Qy	478	PAACVVLSESGKLTKEVODFIAAQVTPTKHLRGVVFVDSIPKGPQTGKLIRKELREIF-	536
Db	481	paavvlehgkmtkeIvdyvasvttakkIrggvfvdevpkgltgkldarkIreili	540
Qy	537	-AORAPKSKL 545	
Db	541	kakkggkskl 550	
RESULT 4			
AAy08523	ID AAY08523 standard; Protein; 550 AA.		
XX	AC AAY08523;		
XX	DT 03-AUG-1999 (first entry)		
XX	DE P. pyralis pTetLucI luciferase protein.		
XX	KW Firefly; luciferase; tetracycline; transcriptional control; TetR; TetA;		
XX	KW tetracycline repressor; tetracycline promoter; luminescence; luxCDABE;		
XX	KW Insect; Tn10; medicine; dosage; cheese production; antibiotic; foodstuff;		
XX	KW allergy.		
XX	OS Photinus pyralis.		
XX	PN WO9925866-A1.		
XX	PD 27-MAY-1999.		
XX	PF 11-NOV-1998; 98WO-FI00873.		
XX	PR 14-NOV-1997; 97FI-0004235.		
XX	PA (KARP/) KARP M.		
XX	PA (KORP/) KORPELA M.		
XX	PA (KURI/) KURITTU J.		
XX	PI Karp M, Korpela M, Kurittu J;		
XX	DR WPI; 1999-338015/28.		
XX	DR N-PSDB; AAV72416.		
XX	PT Assaying for tetracycline using recombinant prokaryotic cells		
XX	PS Example 2; Page 25-26; 67pp; English.		
XX			

This invention describes a novel tetracycline assay that uses recombinant prokaryotic cells comprising a luciferase gene under the transcriptional control of a tetracycline repressor and tetracycline promoter and involves the detection of luminescence emitted from the cells. The assay can be used to distinguish tetracycline from other microbial agents. The invention also describes a novel plasmid comprising either the luxCDABE genes, a tetracycline repressor (TetR) and a tetracycline promoter (tetA) from Tn10, or the insect luciferase gene, a tetracycline repressor (tetR) and a tetracycline promoter (tetA) from Tn10. The tetracycline assay method can be used for the determination of tetracycline in a sample, e.g. to study the dosage and penetration of the medicine. The method can also be used to test cheese production, as cheese making bacteria are not able to work in the presence of tetracycline. The method can also be used to determine the presence or concentration of antibiotics in foodstuffs, e.g. for allergic people. The present assay method does not rely on the growth of microbes as do conventional tests, and so is much more rapid. The present assay is also more sensitive, as even a small amount of luminescence can be detected.

XX Sequence 550 AA;

Query Match 56.3%; Score 1584.5; DB 20: Length 550;  
Best Local Similarity 55.5%; Pred. No. 1.7e-143;  
Matches 305; Conservative 103; Mismatches 137; Indels 5; Gaps 3;

Qy 1 MEE-ENIRHGERPRDIVHPGSAGQOLYOSLYKFASFP--AIIDAHTNEVTSYAQIFETS 57  
Db 1 medaknkkgpapyfpledgtagedqhkamkryalvpgtiaftdahlevnityaeefms 60  
Qy 58 CRLAVSIEOYGLNENNVGVCSENNINFFNPVLAALYIGIPVATSDMYTDELTGHLNI 117  
Db 61 vrleamkryglntnhrivvcenslqfmpvlgalfgavapandynereilnsmml 120  
Qy 118 SKPTMFSSKKALPILRVQOQLSFIRKVVVIDSMYDINGVECVSTFVARYTDHTFDPLS 177  
Db 121 sqptvfvskkgqlklnvqkklplqkllmdsktdyggfsgmyftvtshlppgfneyd 180  
Qy 178 FTKPDFDPLEKIALIMSSSGTGLPKGVVLSHRSLTIRFVHSRDPYIGTRVPQTSILSL 237  
Db 181 fvpesfdrdktialimnssgstgipkvalphrtacvrfshardpifgnqilpdtailsv 240  
Qy 238 VPFHAFGCMFTLSYFVVGKLVVMLKKPEGALFLKTIQNYKIPTIVAPPVVMFLAKSPL 297  
Db 241 vpfhghgmfttlglylfcgfrvmyrfeeeelfrlslqdykqsalvptlfsffakstl 300  
Qy 298 VDQYDLSLITEVATGGAPLGKDVAEAVAKRLKPGIIQGYGLTETCCAVMITPHNAVKTG 357  
Db 301 ldkydlslnhelagsgapskevgeavakrfhipgrrgggigtettsallitpegdkpg 360  
Qy 358 STGRPLPIKARVLONATGKALPGGERGEICFQSSMIMKGYNNPEATIDTIDKDGWLS 417  
Db 361 avgkvvpfeakvvdldtqtlgnqrqelcvggmimsgyynpneactnallidkgwlhs 420  
Qy 418 GDIGYDDEGNFFIVDRLEKILIKYGYQVAPAELENLLQHPSTADAGTVGPDPFEGQOL 477  
Db 421 gdlaywdehffivdrilksilkygyqvapaelesillqhnifdagvaglpddadagel 480  
Qy 478 PAACVVLSESGKTLTEKEVDQFTAAQVTPTKHLRGVGVFVDSIPKQPTCKLRKELREIF- 536  
Db 481 paavvvlhsgkmtelkdvvasqvttakkllrggvvfvdevpvgltgldarkireill 540  
Qy 537 -AQRAPSKL 545  
Db 541 kakkggakl 550

RESULT 5

ID AAB31172  
XX AAB31172 standard; Protein; 550 AA.  
AC AAB31172;  
XX

DT 02-APR-2001 (first entry)  
DE Amino acid sequence of a luciferase protein.  
XX  
KW Growth rate; death rate; reporter gene; luminescent protein;  
KW fluorescent product; luciferase; green fluorescent protein; GFP.  
OS Photinus pyralis.  
XX  
PN WO2000075367-A1.  
XX  
PD 14-DEC-2000.  
XX  
PF 07-JUN-2000; 2000WO-FI00507.  
PP  
PR 07-JUN-1999; 99FI-0001296.  
XX  
PA (LILI/) LILIUS E.  
PA (VIRT/) VIRT M.  
XX  
PI Lilius E, Virta M;  
XX  
DR WPI; 2001-061737/07.  
DR N-PSDB; AAC86954.  
XX  
PT Assessing growth and death rates of a micro-organism in a desired  
PT environment, by introducing 2 reporter genes encoding luminescent and  
XX fluorescent products and detecting luminescent fluorescence -  
XX  
PS Disclosure; Page 27-28; 32pp; English.  
XX  
CC The specification describes a method for assessing the growth rate  
CC and death rate of a micro-organism within a predetermined time period  
CC in a desired environment. The method comprises introducing at least  
CC two reporter genes encoding luminescent and/or fluorescent products  
CC into the micro-organisms, incubating the micro-organism within the  
CC desired environment, and detecting luminescence and/or fluorescence  
CC after a predetermined time period. Use of two different markers within  
CC a micro-organism enables the differentiation between growth and death  
CC rates. The method is used to assess the growth rate and death rate of  
CC a micro-organism within a predetermined time period in a desired  
CC environment. The present sequence represents a luciferase protein,  
CC and is encoded by a plasmid which encodes luminescent and fluorescent  
CC proteins, and is used in the method of the invention.  
XX  
SQ Sequence 550 AA;

Query Match 56.3%; Score 1584.5; DB 22: Length 550;  
Best Local Similarity 55.5%; Pred. No. 1.7e-143;  
Matches 305; Conservative 103; Mismatches 137; Indels 5; Gaps 3;

Qy 1 MEE-ENIRHGERPRDIVHPGSAGQOLYOSLYKFASFP--AIIDAHTNEVTSYAQIFETS 57  
Db 1 medaknkkgpapyfpledgtagedqhkamkryalvpgtiaftdahlevnityaeefms 60  
Qy 58 CRLAVSIEOYGLNENNVGVCSENNINFFNPVLAALYIGIPVATSDMYTDELTGHLNI 117  
Db 61 vrleamkryglntnhrivvcenslqfmpvlgalfgavapandynereilnsmml 120  
Qy 118 SKPTMFSSKKALPILRVQOQLSFIRKVVVIDSMYDINGVECVSTFVARYTDHTFDPLS 177  
Db 121 sqptvfvskkgqlklnvqkklplqkllmdsktdyggfsgmyftvtshlppgfneyd 180  
Qy 178 FTKPDFDPLEKIALIMSSSGTGLPKGVVLSHRSLTIRFVHSRDPYIGTRVPQTSILSL 237  
Db 181 fvpesfdrdktialimnssgstgipkvalphrtacvrfshardpifgnqilpdtailsv 240  
Qy 238 VPFHAFGCMFTLSYFVVGKLVVMLKKPEGALFLKTIQNYKIPTIVAPPVVMFLAKSPL 297  
Db 241 vpfhghgmfttlglylfcgfrvmyrfeeeelfrlslqdykqsalvptlfsffakstl 300  
Qy 298 VDQYDLSLITEVATGGAPLGKDVAEAVAKRLKPGIIQGYGLTETCCAVMITPHNAVKTG 357

```

Db 301 idkydlsnlheiasggaplskevgeavakrfhlpgirggyglttetsailltpgeddkpg 360
Qy 358 STGRPLPYIKAKVLDNATGKALGFCGERGEICFQSEMIMKGYNNPEATIDTIDKDWLHS 417
Db 361 avgkvvpfkaekvvdldgtclgvnqgclcvrgpmimgsvynnppeatnaldkdwls 420
Qy 418 GDICYDDEGCHFFIVDRKLKELIKYGVQVAPAELENLLQHPSTADAGVTGPDEFQGL 477
Db 421 gdiaywdehffivdrklslkygyvapeaesillqhnifdagvaglpddagel 480
Qy 478 PAACVWLESGLTEKEVQDFIAAQVTPTKHLRGVGFVDSIPKGPCKLRKELREIF- 536
Db 481 paavvlehgktmtekeivdyvasqvtakkrlrggvfvdevpvgltgkldarkirelli 540
Qy 537 -AORAPKSKL 545
Db 541 kakkggkskl 550

RESULT 6
AAR83006
ID AAR83006 standard; Protein: 550 AA.
XX
AC AAR83006;
XX
DT 12-FEB-1996 (first entry)
XX
DE Firefly luciferase E354 engineered mutant.
XX
KW Luciferase; thermostability; enzyme stabilization; firefly;
KW Photinus pyralis; luc gene; site-directed mutagenesis; label;
KW ATP analysis.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 354 /note= "amino acid at position 354 is any amino
FT acid other than Glu, pref. also not Gly,
FT Pro or Asp, and is pref. Trp, Val, Leu,
FT Ile, Asn or esp. Arg or Lys, their
FT analogues or modified forms"
XX
XX WO9525798-A1.
XX
XX 28-SEP-1995.
XX
XX 22-MAR-1995; 95WO-GB00629.
XX
XX 20-JAN-1995; 95GB-0001170.
XX 23-MAR-1994; 94GB-0005750.
XX (MINA ) UK SEC FOR DEFENCE.
XX
XX Lowe CR, Murray JAH, Squirrel DJ, White PJ;
XX WPI; 1995-344619/44.
XX N-PSDB; AAT05621.
XX
XX New heat stable luciferase mutants - related DNA, vectors and
XX transformed cells, useful as reporter genes, assay labels and for
XX determinn of ATP
XX
XX Claim 7; Page 20-22; 40pp; English.
XX
XX Novel luciferases are created by site-directed mutagenesis of the
XX P. Pyralis luc gene (see AAT05621) to create substitutions at
XX amino acid position 354 (glutamate in the wild-type) of the encoded
XX protein (AAR83006) that result in improved enzyme thermostability.
XX
XX Sequence 550 AA;

```

```

Query Match 56.38; Score 1583.5; DB 16; Length 550;
Best Local Similarity 55.5%; Pred. No. 2.le-143;
Matches 305; Conservative 103; Mismatches 137; Indels 5; Gaps 3;

Qy 1 MEE-ENIRHGERPRDIVHPGSAGQOLYOSLYKFKASFP--ALIIDAHTNEVTSYAQIFETS 57
Db 1 medaknkikgkppfypiedgtageqlhkmxryalvpgtiaftahleavniyaeyfems 60
Qy 58 CRLAVSTBOYGLNENNVGVCSENNINFPVLAALYLGIPVATSNMDYTGELTGHINI 117
Db 61 vrleaeamryglntnhrivvcensnlqffmpvlgalfigvavapendiynereellnsmni 120
Qy 118 SKPRIMSSKALPILRVQONLSFIKAVVWIDSWYDINGECVSTFVARVTDHFDPLS 177
Db 121 sqptvvfvskgqlknvqkklplqkllimdsksdyggfsgmyftvtshlppgfneyd 180
Qy 178 FTPKDFDPLEKIALIMSSSGTTGLPKGVVLSHRSLTIRFVHSRDPYIGTRVPOTSILSL 237
Db 181 fypesfdrdkktialimssgstgipkgvalphrtacvrfshardpifgnqilpdcailsv 240
Qy 238 VPFHAFGMFTTLSYFVVGLKVVMCLKFEGALFLKTIQNYKIPTIVVAPPVVMVFLAKSPL 297
Db 241 vpfhhgfmfttlylgicfrvvlmyrfeeeelfrlsldqykqlsallvptlfsfakstl 300
Qy 298 VDQYDLSLSEVATGGAPLGKDVAAVAARLKLPGIIOGYGLTETCCAVMTTPHNAVKTG 357
Db 301 idkydlsnlheiasggaplskevgeavakrfhlpgirggyglttetsailltpxgddkpg 360
Qy 358 STGRPLPYIKAKVLDNATGKALGFCGERGEICFQSEMIMKGYNNPEATIDTIDKDWLHS 417
Db 361 avgkvvpfkaekvvdldgtclgvnqgclcvrgpmimgsvynnppeatnaldkdwls 420
Qy 418 GDICYDDEGCHFFIVDRKLKELIKYGVQVAPAELENLLQHPSTADAGVTGPDEFQGL 477
Db 421 gdiaywdehffivdrklslkygyvapeaesillqhnifdagvaglpddagel 480
Qy 478 PAACVWLESGLTEKEVQDFIAAQVTPTKHLRGVGFVDSIPKGPCKLRKELREIF- 536
Db 481 paavvlehgktmtekeivdyvasqvtakkrlrggvfvdevpvgltgkldarkirelli 540
Qy 537 -AORAPKSKL 545
Db 541 kakkggkskl 550

RESULT 7
AAR98518
ID AAR98518 standard; Protein: 550 AA.
XX
AC AAR98518;
XX
DT 11-MAR-1997 (first entry)
XX
DE P. pyralis variant luciferase E270K/E354K.
XX
KW Luciferase; firefly; ATP; heat stability; luciferin; light.
XX Photinus pyralis.
XX
XX Key Location/Qualifiers
XX FT Misc-difference 270
XX FT /label= E270K
XX FT Misc-difference 354
XX FT /label= E354K
XX
XX WO9622376-A1.
XX
XX 25-JUL-1996.
XX
XX 19-JAN-1996; 96WO-GB000099.
XX
XX 24-APR-1995; 95GB-0008301.

```

```
PR 20-JAN-1995; 95GB-0001172.
XX (MINA ) UK SEC FOR DEFENCE.
XX
XX Lowe CR, Murray JA, Squirrell DJ, White PJ;
XX WPI; 1996-354533/35.
XX
XX Mutant luciferase(s) with greater heat stability than wild-types -
XX useful in luminescent reagents
XX
XX Example 2; Page 26-29; 39pp; English.
XX
XX This sequence represents a variant luciferase from the firefly, P.
XX pyralis. In this sequence the amino acid at position 270 has
XX been mutated such that it is not Glu. Also the amino acid at position
XX 354 has been mutated so that it is not Glu. This luciferase has a
XX lower Km for the substrate ATP than native luciferase. This
XX luciferase therefore has a greater heat stability than wild-type
XX luciferase. The luciferase may be used in an assay to measure ATP
XX with luciferin, whereby the amount of light is related to the amount
XX of ATP present.
XX
XX Sequence 550 AA;
SQ
Query Match 56.3%; Score 1583.5; DB 17; Length 550;
Best Local Similarity 55.5%; Pred. No. 2.le-143;
Matches 305; Conservative 103; Mismatches 137; Indels 5; Gaps 3;
Qy 1 MEE-ENIRHGERPRDIVHPGSAGQOOLYOSLYKFAFPE--AIIDAHTNEVISYAOIFETS 57
Db 1 medaknkkpappfyldgedgagelqhkamkryalvpgtiaftdahlevnityaeefems 60
Qy 58 CRLAVSIEQYGLNENNVGVCSENNINFPNLAALYIGIPVATSNMDYTGELTGHINI 117
Db 61 vrleamkryglnthnrvicvnsnslqfmpvlgalfgavavapandynereilnsmni 120
Qy 118 SKPTIMSSKKALPLILRVQONLSFIKKVVIDSMYDINGVECVSTFVARVTDHTFDPLS 177
Db 121 sqptvfvskkgkqlnkvqkklpiqkllmdsktdyggfsgmyftvtshlpqgfneyd 180
Qy 178 FTPKDFDPLEKIALIMSSSGTGLPKGVLSHRSLTIRFVHSRDPDIYGRTRVPOTSILSL 237
Db 181 fvpesfrdrktialimsssgtgipkpgvalphrtacvrfshardpifgnqilpdtailsv 240
Qy 238 VPFHAFGCMFTLSYFVVGLKVMKKEGALFLKTIQNYKIPTIVVAPPVWVFLAKSPL 297
Db 241 vpfhngfmgfttlylcygfrvmyrfekeflrsldqykgallvptlfsffakstl 300
Qy 298 VDQYDLSLSEVATGGAPLGKDVAAEAVAKRLKPGIIOGYGLTETCCAVMITPHNAVKTG 357
Db 301 idkydlnlhelasgapsksvegcaavakrhlpgirggyglttetsailtpkqddkpg 360
Qy 358 STGRPLPIYAKVLDNATKALGPERGEICFQSEIMMKGYNNPEATIDIKDGLWHS 417
Db 361 avgkvvpfreakvddlgtklgnvqrgelevrgpmimgvynppeatnalldkdgwlhs 420
Qy 418 GDIGYDDEGNFFIVDRLEKILIKYGYQVAPAELENLLOHPSTADACVGTVPDEFQGL 477
Db 421 gdiaywdehffivdrllksllkygyqvapaelesillqbpnifdagvagipddagel 480
Qy 478 PAACVVLSESGKLTKEVQDFIAAQVTPTKHLRGGVVFVDSIPKPTCKLIRKELREIF- 536
Db 481 paavvvlehgtmtkeleivdyasqvttakkrlggvvfdevpvgltgklarkirelli 540
Qy 537 -AQRAPSKL 545
Db 541 kakkggkskl 550
```

RESULT 8  
AAB35498

```
ID AAB35498 standard; Protein; 552 AA.
XX
XX AAB35498;
XX
XX 10-DEC-2001 (first entry)
XX
XX Synthetic luciferase.
XX
XX Synthetic luciferase; thermostability; wavelength shift;
XX bioluminescent assay.
XX
XX Synthetic.
XX
XX WO200131028-A2.
XX
XX 03-MAY-2001.
XX
XX 26-OCT-2000; 2000WO-GB04133.
XX
XX 26-OCT-1999; 99GB-0025161.
XX
XX 10-JUL-2000; 2000GB-0016744.
XX
XX (MINA ) UK SEC FOR DEFENCE.
XX
XX Squirrell DJ, Murphy MJ, Price RL, White PJ, Willey TL;
XX
XX WPI; 2001-308651/32.
XX
XX N-PSDB; AAF85716.
XX
XX Novel mutant luciferase for bioluminescent assays, determining presence
XX of COA in a sample for diagnosing diabetes, has enhanced
XX thermostability and emits light at different wavelength compared to
XX wild-type luciferase -
XX
XX Example 12; Fig 14; 81pp; English.
XX
XX The present invention describes the synthetic protein and coding
XX sequences of a luciferase enzyme with increased thermostability and a
XX shift in wavelength compared to natural luciferases. The protein can be
XX used in bioluminescent assays using the luciferase/luciferin reaction as
XX a means of signalling, particularly for determining the presence of COA
XX in a sample for use in the diagnosis of diabetes. The present sequence is
XX the protein of the invention.
XX
XX Sequence 552 AA;
Qy
Query Match 56.3%; Score 1583.5; DB 22; Length 552;
Best Local Similarity 55.5%; Pred. No. 2.le-143;
Matches 305; Conservative 102; Mismatches 138; Indels 5; Gaps 3;
Qy 1 MEE-ENIRHGERPRDIVHPGSAGQOOLYOSLYKFAFPE--AIIDAHTNEVISYAOIFETS 57
Dy 3 medaknkkpappfyldgedgagelqhkamkryalvpgtiaftdahlevnityaeefems 62
Qy 58 CRLAVSIEQYGLNENNVGVCSENNINFPNLAALYIGIPVATSNMDYTGELTGHINI 117
Dy 63 vrleamkryglnthnrvicvnsnslqfmpvlgalfgavavapandynereilnsmni 122
Qy 118 SKPTIMSSKKALPLILRVQONLSFIKKVVIDSMYDINGVECVSTFVARVTDHTFDPLS 177
Dy 123 sqptvfvskkgkqlnkvqkklpiqkllmdsktdyggfsgmyftvtshlpqgfneyd 182
Qy 178 FTPKDFDPLEKIALIMSSSGTGLPKGVLSHRSLTIRFVHSRDPDIYGRTRVPOTSILSL 237
Dy 183 fvpesfrdrktialimsssgtgipkpgvalphrtacvrfshardpifgnqilpdtailsv 242
Qy 238 VPFHAFGCMFTLSYFVVGLKVMKKEGALFLKTIQNYKIPTIVVAPPVWVFLAKSPL 297
Dy 243 vpfhngfmgfttlylcygfrvmyrfekeflrsldqykgallvptlfsffakstl 302
Qy 298 VDQYDLSLSEVATGGAPLGKDVAAEAVAKRLKPGIIOGYGLTETCCAVMITPHNAVKTG 357
Dy 301 idkydlnlhelasgapsksvegcaavakrhlpgirggyglttetsailtpkqddkpg 360
```

Db	303	idkydlnlnheiasgapsksvegavakrflhplgrqgygltettsailitcpkdgfkpg	362
Qy	358	STGRPLPYIKAKVLDMATGKALGPGRGETICFOSEMIMKGYNNPEATIDTIDKDGWLHS	417
Db	363	avgkvpvffeakvvdldtktlgvnrqgelcvgpmimgvynpeatnalidkdgwlhs	422
Qy	418	GDIGYVDEGDNFFIVDLRLKELIKYGYVAPAELENLLLOHPSIADAGVTGVPDEFGGOL	477
Db	423	gdlaywdehffivdrllsklkygygvapaelesillqhpnifdagvaglpdddagel	482
Qy	478	PAACVWLESGKTUTEKEVODFIAAQTPTPKHLRGVVFVDSIPKGTGKLIRKELREIF-	536
Db	483	paavvlehgktkteikdyvasqvtakkirggvfvdevpklgtgldarkirelli	542
Qy	537	-AQRAPKSKL 545	
Db	543	kakgggkskl 552	
RESULT 9			
ID	AAR28127	standard; Protein; 815 AA.	
XX	AAR28127;		
XX			
DT	12-MAR-1993	(first entry)	
XX			
DE	Lux::npt-II fusion protein.		
XX			
KW	Neomycin phosphotransferase-II; genetic selection; genetic marker;		
KW	gene recovery.		
XX			
OS	Synthetic.		
XX			
PN	WO9217593-A.		
XX			
PD	15-OCT-1992.		
XX			
PF	30-MAR-1992; 92WO-CA00139.		
XX			
PR	28-MAR-1991; 91US-0676432.		
XX			
PA	(CANADA) NAT RES COUNCIL CANADA.		
XX			
PI	Crosby WL, Datla RSS, Hammerlindl JK, Selvaraj G;		
XX			
DR	WPI; 1992-366262/44.		
DR	N-PSDB; AAQ30000.		
XX			
PT	Fused gene with characteristics of component parts - comprising a		
PT	gene conferring conditional growth advantage and a marker gene,		
PT	used as a probe to select and isolate genetic elements		
XX			
PS	Disclosure; Page 33; 53pp; English.		
XX			
CC	The sequence is that of the fusion protein encoded by the firefly		
CC	{Pyrallis} luciferase:neomycin phosphotransferase-II chimeric gene.		
CC	See also AAR28125.		
XX			
SQ	Sequence 815 AA;		

Query Match	56.2%	Score 1581.5;	DB 13;	Length 815;
Best Local Similarity	55.5%;	Pred. No. 6.2e-143;		
Matches	305;	Conservative 103;	Mismatches 137;	Indels 5; Gaps

  

Qy	1	MEE--ENTRHGERPRDIVHPSGAGQLYQSILYKFPASPE--AIIADHTNEVSVYQAIFETS	57
		:    :    :    :    :    :    :    :    :    :    :    :    :	
Db	1	medaknikkppapfpiedgtageihkamkryalypgtlaftadhaievnityaefems	60
Qy	58	CRLAVSTEQYGLNENNVVGVCSENNINFFNPVLAAIYLGIPVATSDNDMTDGLTGLHLM	117
		:    :    :    :    :    :    :    :    :    :    :    :	
Db	61	vrltaeamkryglntnrhivvcvssnslqffmpvlgalfigvavapandiynerellgysmli	120

Qy	118	SKTIPMSKKALPLILRVOQNLSFTKVVVVVDSMDYDINCVCVSTFVARYTDHTFDPLS	177
Db	121	sqptvfvsvkkgqlilnvqkklpilqkiiimsktdyqfgsmylvfvtshippgfnedy	180
Qy	178	FTPKDQDPLKIALIMSSSGTGLPKGVWLSHRSLSIRFVHSRDPITYGRTVPOTSLSL	237
Db	181	fvpsfardktialimsssgstglpkvalphrtacvrfshardpivgnqilpdtailsv	240
Qy	238	VPHHAFGMFTTISYFVVGUGLKVWMLKKKFFEGALPLKTIQNYKTIIVVAVPVWVFLAKSPL	297
Db	241	vpfnhvgfmltglylicgrfvrlmyrfeseellfrsqdykqslallvptlfsfakstl	300
Qy	298	VDQYDLSLSLEVATGGAPLCKDVAEAVAKRLKLPJIGQYGLTETCCAVMTPHNAVKTG	357
Db	301	idkydlnllheiasggaplskevgaeavakrfhlpgrirqyglletcsailitpegddkpg	360
Qy	358	STGRPLPYIAKVLNDNATGKALPGERGEICFOSEIMKGYNNPEATIDTIDKDGWLHS	417
Db	361	avgkvpvfeakvvdltgktlgnvqrgelcivrpgmlmsgyvnnpeatnalldkdgwllhs	420
Qy	418	GDICYDEDEGNFFTVDRLLKELIKYGVQVAPAELENLLQHPISADAGVTGVDPDEGGOL	477
Db	421	gdlaywedeheffivdrkslikykygvapaelesilhqhpnifdagvagipddagel	480
Qy	478	PAACVWVLESKTLTEKEVQODFIAAQVTPPTKHLRGVVVFVDSIPKGBTGLKRLKEIRF-	536
Db	481	paavvviehgktmtekeivdyvasqvtakkllrgvvfvdevpkgtlglkldarkirefll	540
Qy	537	-AQRAPKSKL 545	
Db	541	kakkggkskl 550	

RESULT 10  
AAR54867  
ID AAR54867 standard; Protein; 743 AA.

DT 21-DEC-1994 (first entry)

DE Hepatitis C Virus core protein-luciferase fusion protein.  $\Delta\Delta$ 

Hepatitis C Virus; HCV; control; Non-A, non-B hepatitis virus;  
antiserum; antisense; therapy; inhibition; viral protein precursor;  
recombinant vaccinia virus; HCV core protein gene; firefly;  
luciferase reporter gene; fusion construct.

AA	Chimeric Hepatitis C virus.
OS	Chimeric Photinus pyralis.
OS	Chimeric Photinus pyralis.

CA2104649-A.

26-FEB-1994.

XX 23-AUG-1993: 93CA-2104649

XX  
PB 25-AUC-1092. 92PB-0248796

PR 03-MAR-1993; 93JP-0042736.  
yy

PA (SEKI//) SEKI M.

PI Honda Y, Seki M, Yamada E;

DR WPI; 1994-151836/19.

XX

PT virus genome - are useful as

PS Example 5; Page 244-250; 262

CC A recombinant vaccinia virus which codes for a HCV core protein  
CC fused to the firefly luciferase enzyme was constructed from PCR  
CC amplified fragments. The construct (which encodes the fusion  
CC protein sequence AAR54867) was useful for assaying the inhibitory  
CC activity of various antisense oligonucleotides on HCV gene  
CC translation; detection of luciferase activity provided a measure  
CC of expression of the HCV core protein.  
XX  
SQ Sequence 743 AA;

Query Match 56.1%; Score 1577.5; DB 15; Length 743;  
Best Local Similarity 55.3%; Pred. No. 1.3e-142;  
Matches 304; Conservative 103; Mismatches 138; Indels 5; Gaps 3;  
QY 1 MEE-ENIRHGERPRDIVHPGSAGQOLYOSLYKFASPE--AIIAHTNEVISYAOIFETS 57  
DB 194 medaknlkkgpapfpyledgtagedlhamkryalvpgtiaftdhdhievnityaeefms 253  
QY 58 CRLAVSIEQYGLNENNVGVCSENNFNPVLAALYLIGIPVATSDMYDGTGELTGHLLNI 117  
DB 254 vrlaamkrygintnhrivcsenslqfmpvlgalfgavavapandynereellnsmni 313  
QY 118 SKPTTFSSKKALPILRVQNLSPKIKVVDSDMYDINGECVSTFVARYTDHTFDPLS 177  
DB 314 sqptvfvskkgqlklnvqkklpiqkiiimdsdtdyggfsgmyftvshlppgfneyd 373  
QY 178 FPKDFDPLEKIALTMSSSGTGLPKGVVLSHRSUTIRFVHSRDIYGTTRVPTQSILSL 237  
DB 374 fvpesfdrdktiallmnsagstgpkgvvalphrtacvrfshardpifgnlqldcallav 433  
QY 238 VPFHAFGMFTTSLYFVVGKLVMLKKEGALFLKTIQNIKIPTIVVAPPVNVFLAKSPL 297  
DB 434 vpfhgmfttlylgicfrvvlmyrfeeeelfrslqklsallvptlfsffakstl 493  
QY 298 VDOYDLSLSEVATGAPLGKDVAEAVAKRLKLPGLIOGYLGTETCCAVMTPHNAVKTG 357  
DB 494 ldkydlslheliasgagplskvegeavakrfhlpgirggygtettsallitpegddkpg 553  
QY 358 STGRPLPYIKARVLNATCKALGPGEIGEICFQSEIMKMGYNNPNEATIDTIDKDWLHS 417  
DB 554 avgvvpfpeakvvdldgtklgnvqrgelcvrgpmimgsvyvnneatnallidkgwlhs 613  
QY 418 GDIGYDEDEGNFFIVDRKLKELIKYGYQVAPAELENLLQHPISADAGVTGVPDEFEGOL 477  
DB 614 gdiaywdedehffivdlklslikygyqvapaelesillqhnifdagvaglpdddagel 673  
QY 478 PAACVLESKGTLTEKEVQDFTAAQVTPTKHLRGGVGVFVDSIPKGTCKLRKELREIF- 536  
DB 674 paavvvhgkmtktekeivdyvasqvtctakkirggyvfvdevpklgtkldarkirelli 733  
QY 537 -AQRAPKSKL 545  
DB 734 kakkggkkl 743

RESULT 11  
ID AAR72801  
XX AAR72801 standard; Protein; 743 AA.  
AC AAR72801;  
DT 04-DEC-1995 (first entry)  
DE Vaccinia virus vector comprising HCV T7N1-19 and firefly luciferase.  
XX Vaccinia virus vector; firefly luciferase; hepatitis C virus; T7N1-19;  
KW antiviral agent; poliovirus; human rhinovirus;  
KW internal ribozyme entry site; non-A non-B; cerebral cardio-hepatitis;  
KW foot and mouth disease.  
XX Synthetic.  
OS  
XX

PN JP07069899-A.  
XX 14-MAR-1995.  
XX 02-SEP-1993; 93JP-0241973.  
XX 02-SEP-1993; 93JP-0241973.  
XX (MITU ) MITSUBISHI KASEI CORP.  
XX WPI; 1995-144713/19.  
DR N-PSDB; AAR86799.  
XX Antiviral agent comprising component which disrupts viral gene  
PT translation - used for the selective inhibition of e.g. Hepatitis  
PT C virus, polio:virus and human rhinovirus  
XX Disclosure; Pages 19-22; 23pp; Japanese.  
XX AAR86788 encodes AAR72800 Hepatitis C virus (HCV) T7N1-19, which  
CC disrupts viral gene translation, by preventing the binding of  
CC the viral mRNA to the internal ribozyme entry site. It was  
CC used in the construction of an antiviral agent detecting vector  
CC (AAR86799 which encodes AAR72801), which comprises a vaccinia virus  
CC vector, a vector containing T7N1-19 and a firefly luciferase gene.  
CC The antiviral agent can be used to treat HCV, poliovirus, cerebral  
CC cardio-hepatitis, human rhinovirus and foot and mouth disease viral  
CC infections.  
XX Sequence 743 AA;  
SQ  
Query Match 56.1%; Score 1577.5; DB 16; Length 743;  
Best Local Similarity 55.3%; Pred. No. 1.3e-142;  
Matches 304; Conservative 103; Mismatches 138; Indels 5; Gaps 3;  
QY 1 MEE-ENIRHGERPRDIVHPGSAGQOLYOSLYKFASPE--AIIAHTNEVISYAOIFETS 57  
DB 194 medaknlkkgpapfpyledgtagedlhamkryalvpgtiaftdhdhievnityaeefms 253  
QY 58 CRLAVSIEQYGLNENNVGVCSENNFNPVLAALYLIGIPVATSDMYDGTGELTGHLLNI 117  
DB 254 vrlaamkrygintnhrivcsenslqfmpvlgalfgavavapandynereellnsmni 313  
QY 118 SKPTTFSSKKALPILRVQNLSPKIKVVDSDMYDINGECVSTFVARYTDHTFDPLS 177  
DB 314 sqptvfvskkgqlklnvqkklpiqkiiimdsdtdyggfsgmyftvshlppgfneyd 373  
QY 178 FPKDFDPLEKIALTMSSSGTGLPKGVVLSHRSUTIRFVHSRDIYGTTRVPTQSILSL 237  
DB 374 fvpesfdrdktiallmnsagstgpkgvvalphrtacvrfshardpifgnlqldcallav 433  
QY 238 VPFHAFGMFTTSLYFVVGKLVMLKKEGALFLKTIQNIKIPTIVVAPPVNVFLAKSPL 297  
DB 434 vpfhgmfttlylgicfrvvlmyrfeeeelfrslqklsallvptlfsffakstl 493  
QY 298 VDOYDLSLSEVATGAPLGKDVAEAVAKRLKLPGLIOGYLGTETCCAVMTPHNAVKTG 357  
DB 494 ldkydlslheliasgagplskvegeavakrfhlpgirggygtettsallitpegddkpg 553  
QY 358 STGRPLPYIKARVLNATCKALGPGEIGEICFQSEIMKMGYNNPNEATIDTIDKDWLHS 417  
DB 554 avgvvpfpeakvvdldgtklgnvqrgelcvrgpmimgsvyvnneatnallidkgwlhs 613  
QY 418 GDIGYDEDEGNFFIVDRKLKELIKYGYQVAPAELENLLQHPISADAGVTGVPDEFEGOL 477  
DB 614 gdiaywdedehffivdlklslikygyqvapaelesillqhnifdagvaglpdddagel 673  
QY 478 PAACVLESKGTLTEKEVQDFTAAQVTPTKHLRGGVGVFVDSIPKGTCKLRKELREIF- 536  
DB 674 paavvvhgkmtktekeivdyvasqvtctakkirggyvfvdevpklgtkldarkirelli 733  
QY 537 -AQRAPKSKL 545



```
XX CC The present sequence is encoded by the modified firefly luciferase
CC CDNA luc+, which can be used as a reporter gene in experimental
CC biological systems, e.g. cell cultures, transgenic plants and
CC animals, and in cell free expression systems. Removal of the CDNA's
CC peroxisomal translocation sequence yields a cytoplasmic form of
CC the enzyme with optimal substrate availability. Potentially
CC interfering restriction sites and genetic regulatory sites are
CC removed, and codon usage for mammalian cells is improved. Absence
CC of potential N-glycosylation sites minimises post-translational
CC modification.
XX SQ Sequence 550 AA;

Query Match 55.8%; Score 1568.5; DB 18; Length 550;
Best Local Similarity 54.7%; Pred. No. 5.9e-142;
Matches 300; Conservative 103; Mismatches 142; Indels 3; Gaps 2;

Qy 1 MEE-ENIRGERPRDIVHPGAGQOOLYQSLYKFAFPE--AIIDAHTEVISAQIFETS 57
Db 1 medaknikkppapfpledgtageqlhkamkryalvpgtiaftdahievdtiaeyfems 60

Qy 58 CRLAVSIEQYGLNENNVGVCSENNINFPNVLALYLGIPVATNDMYTDCGELTGH LNI 117
Db 61 vrleamkrygintnhrivcsenslqfmpvlgalfgavavapandiynerellnsmgi 120

Qy 118 SKPTIMFSSKALPLILRVQOQLSPFKKVVVIDSMYDINGVECVSTFVARVTDHTFDPLS 177
Db 121 sqptvvfsvsk9qlklnvqkklpiqlkllmdsktdyggfsgmyfvtshlppgfneyd 180

Qy 178 FTKDPDPLEKIALIMSSSGTGLPKGVLSHRSLTIRFVHSRDPYIGTRVTPQTSILSL 237
Db 181 fvpsfdrdktalimnssgstglpkgvallphrtacvrfshardpifgnqilpdtailsv 240

Qy 238 VPFHAFGCMFTTLSYFVGLKVVMLKKFEGALFLKTIQNYKIPTIVVAPPVWVFLAKSPL 297
Db 241 vpfhgfmgfttlgylcgfrvvlmyrfeelflrslqdyklsallvptlfsfakatl 300

Qy 298 VDOYDLSLSEVATGCGAPLGDVAEAVAKRLKLPGLIOGYGLTETCCAVMITPHNAVKTG 357
Db 301 idkdylnlhelasgagaplskegavakrfhlpgirggygltectsailltpegddkpg 360

Qy 358 STGRPLPYIKARVLNATGKALGPGERGEICFQSEIMKGYNNPEATIDTDKDGWLHS 417
Db 361 avgvkvpfteeskvdldgktlgnvrgelcvgmimgynnppeatnalldkgwlhs 420

Qy 418 GDIGYDEGDNFFIVDRKLKELIKYGYOVAEPALENLLQHPSIADAGVTGVPDEFEGQOL 477
Db 421 gdlaywdedehffivdrilkslikygyqvapaelesillqhnifdagvaglpdddagel 480

Qy 478 PAACVVLESGKLTLEKEVQDFAAQTPTKHLRGVGVFVDSIPKQPTCKLRKELREIFA 537
Db 481 paaavvlehngkmteteivdyvasqvtcakkirgvgvfvdevpklgtldkarkirelli 540

Qy 538 QRAPKSKL 545
Db 541 kakkggki 548

RESULT 14
ID AAB35055
XX AAB35055 standard; Protein; 550 AA.
XX AC AAB35055;
XX 27-MAR-2001 (first entry)
XX Plant signal transduction pathway reporter vector pZP-ALE protein.
XX Plant signal transduction pathway; reporter vector;
XX luciferase; herbicide resistance; pest resistance.
```

```
OS Synthetic.
XX WO200071668-A2.
XX 30-NOV-2000.
XX 22-MAY-2000; 2000WO-US14041.
XX 20-MAY-1999; 99US-0136145.
XX (UYNE-) UNIV NEBRASKA.
XX Rhoades DW;
XX WPI; 2001-025143/03.
XX N-PSDB; AAC66931.
XX Identifying genes encoding signal transduction components useful for
XX producing transgenic plants, by transforming plants with vector
XX encoding reporter gene, mutating the plant and identifying genes from
XX the mutant -
XX Disclosure; Page 47-49; 54pp; English.
XX The present invention provides a method for identifying genes encoding
XX components of plant signal transduction pathways between mitochondrial
XX function and metabolic status and nuclear gene expression. This involves
XX the transformation of a plant with a vector encoding a reporter gene
XX linked to an AOX promoter, mutagenesis of the plant to increase
XX expression of the reporter gene and then determining the identity of the
XX signal transduction gene of interest. This is useful in the production of
XX transgenic plants with increased productivity, herbicide, stress and pest
XX resistance.
XX Sequence 550 AA;

Query Match 55.8%; Score 1568.5; DB 22; Length 550;
Best Local Similarity 54.7%; Pred. No. 5.9e-142;
Matches 300; Conservative 103; Mismatches 142; Indels 3; Gaps 2;

Qy 1 MEE-ENIRGERPRDIVHPGAGQOOLYQSLYKFAFPE--AIIDAHTEVISAQIFETS 57
Db 1 medaknikkppapfpledgtageqlhkamkryalvpgtiaftdahievdtiaeyfems 60

Qy 58 CRLAVSIEQYGLNENNVGVCSENNINFPNVLALYLGIPVATNDMYTDCGELTGH LNI 117
Db 61 vrleamkrygintnhrivcsenslqfmpvlgalfgavavapandiynerellnsmgi 120

Qy 118 SKPTIMFSSKALPLILRVQOQLSPFKKVVVIDSMYDINGVECVSTFVARVTDHTFDPLS 177
Db 121 sqptvvfsvsk9qlklnvqkklpiqlkllmdsktdyggfsgmyfvtshlppgfneyd 180

Qy 178 FTKDPDPLEKIALIMSSSGTGLPKGVLSHRSLTIRFVHSRDPYIGTRVTPQTSILSL 237
Db 181 fvpsfdrdktalimnssgstglpkgvallphrtacvrfshardpifgnqilpdtailsv 240

Qy 238 VPFHAFGCMFTTLSYFVGLKVVMLKKFEGALFLKTIQNYKIPTIVVAPPVWVFLAKSPL 297
Db 241 vpfhgfmgfttlgylcgfrvvlmyrfeelflrslqdyklsallvptlfsfakatl 300

Qy 298 VDOYDLSLSEVATGCGAPLGDVAEAVAKRLKLPGLIOGYGLTETCCAVMITPHNAVKTG 357
Db 301 idkdylnlhelasgagaplskegavakrfhlpgirggygltectsailltpegddkpg 360

Qy 358 STGRPLPYIKARVLNATGKALGPGERGEICFQSEIMKGYNNPEATIDTDKDGWLHS 417
Db 361 avgvkvpfteeskvdldgktlgnvrgelcvgmimgynnppeatnalldkgwlhs 420

Qy 418 GDIGYDEGDNFFIVDRKLKELIKYGYOVAEPALENLLQHPSIADAGVTGVPDEFEGQOL 477
Db 421 gdlaywdedehffivdrilkslikygyqvapaelesillqhnifdagvaglpdddagel 480
```





**This Page Blank (uspto)**





GENERAL INFORMATION:  
APPLICANT: SQUIRRELL, DAVID J.  
APPLICANT: LOWE, CHRISTOPHER R.  
APPLICANT: WHITE, PETER J.  
APPLICANT: MURRAY, JAMES A. H.  
TITLE OF INVENTION: MUTANT LUCIFERASES  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: NIXON & VANDERHUYE P.C.  
STREET: 1100 NORTH GLEBE ROAD  
CITY: ARLINGTON  
STATE: VIRGINIA  
COUNTRY: U.S.A.  
ZIP: 22201-4714  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/875,277A  
FILING DATE: 01-OCT-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: GB 9501172-2  
FILING DATE: 20-JAN-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9508301.0  
FILING DATE: 24-APR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: CRAWFORD, ARTHUR R.  
REGISTRATION NUMBER: 25,327  
REFERENCE/DOCKET NUMBER: 124-588  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 816-4000  
TELEFAX: (703) 816-4100  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 550 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: Photinus pyralis  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 270  
-08-875-277A-2

[illegible]

```

Db      241  ||||| ||||| | : | :||: || ||||::||| : :: | : | ||| |
          VPFHGGFMFTTGLGYLCGRVVMYRFEXELFRLSLQDKIQSALLVPTLFSFFAKSTL 300

Qy      298  VDQYDLSLSEVATGGAPLGKDAEAVAKRLKLPGIIIGQYGLTETCCAYMITHNAVKTG 357
          :|: |:||||| | :| ||||| :| ||||| ||||| | :||| | |
Db      301  IDKTDJLNHLHEISGGAPLSKEVEAVAKRPHLPGINQGYYGLTETSAILITPEGDCKPG 360
          | :|: |:||||| | :| ||||| :| ||||| ||||| | :||| | |
Qy      358  STGRPLPYIKAKVLNDNATKALGPGERGEICFOSEMIMKGYYNNPEATIDTIDKDGWLHS 417
          :|: |:||||| ||| | :| :||:-| ||| ||||| ||||| | :||| | |
Db      361  AVGVVPFEFAKVVDLDGTGLGVNQRGELCVRPMMSGVYNPEATNALIDKDGWLHS 420
          :|: |:||||| ||| | :| :||:-| ||| ||||| ||||| | :||| | |
Qy      418  GDYGYYDEBGNFFIVDRLELIKKYQVAPAELENLLQHPSIADAGVTGVPDEFEGQL 477
          ||| | :|: |:||||| ||||| ||||| ||||| ::|||: ||| | :||: | :|
Db      421  GDIATWDEDEHFIVDRLEKSLIKYQVAPAELESILLQHPNIFDQVAGLPDDDAGEL 480
          ||| ||||| DEDEHFIVDRLEKSLIKYQVAPAELESILLQHPNIFDQVAGLPDDDAGEL 480

Qy      478  PAACVWLESCKTTTEKEVQDFIAAQTTPTKHLRGVVVFVDSIPKPGTGKILIRKELREIF- 536
          ||| ||||| LESCKTTTEKEVQDFIAAQTTPTKHLRGVVVFVDSIPKPGTGKILIRKELREIF- 536
Db      481  PAAVVLHEHGKTMTEKEIVDIVASQVTTAKKLRGGVVVFDEVEPKGLTGKLDARKIRILI 540
          ||| ||||| HEHGKTMTEKEIVDIVASQVTTAKKLRGGVVVFDEVEPKGLTGKLDARKIRILI 540

Qy      537  -AQRAPKSKL 545
          |::: ||||

Db      541  KAKKGGSKL 550
          |||||

RESULT 4
US-08-867-352-23
; Sequence 23, Application US/08867352
; Patent No. 6060273
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Multicistronic expression units and their use
; NUMBER OF SEQUENCES: 25
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (BPA)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/867,352
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/387,847
; FILING DATE:
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 550 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SS-08-867-352-23

```

Query Match	56.3%	Score 1584.5;	DB 3;	Length 550;
Best Local Similarity	55.5%;	Pred. No. 1.4e-153;		
Matches 305; Conservative 103; Mismatches 137; Indels 5; Gaps 3;				
y	1	MEE-ENIRGGERPRDITVPHGSAGQQLYSKFASPE--AIIDAHTNEVISYAOIFETS	57	
b	1	MEDAKNIKKGPAPPYPLEDTAGEQLHKMKRYALVPGTIAFTDAHIEVNITYAEVFEMS	60	
y	58	CRLAVSIEOGLMENNVMVGVCSENNTNFNPVALYLGIPIVATSNMDYTDELGTGLHNI	117	
b	61	VRLAEAMKRYGLTNHRIVVCSSENSIQFFMPVLGALFIGVAPANDIYNEREELLNSMI	120	
y	118	SKEPTIFFSKKALLPLIRLQQONLSFTKKVVVIDSMVDINGECVSTEVARYTDHRTFDPLS	177	
b	121	SQTFFVFSKKGLOKLINVKOKLPLIOKIIIMDSKTDYGCFOSMTFVTYSHLPFGNEYD	180	
y	178	FTPDKDFPLEKTIALIMSSSGTTGLPKGVVLSHSLTIRFVHSRDPRIYGRTRTPQTSILSL	237	
b	181	FVPSFEDRKOTKIALIMNNSSGTGLPKGVALPHTACVRESHARDPIFGNQIIPDTAILSV	240	

QY 238 VPFHAFGMFTTSLSYFVGLKVVMLKKFEGALFLKTIQNIKIPTIVAPPVVMFLAKSPL 297  
Db 241 VPFHGFGMFTTGLYLICGFRVLMYRFEELFLASLDQYKIQSALLVPTLFSFFAKSTL 300  
QY 298 VDQYDLSLSEVATGAGPLGKDVAAVAKRLKPLGLIOGYGLTETCCAVMTPHNAVKTG 357  
Db 301 IDKYDLSNLHEIASGAPLSKEVGAARFHLPGIRQGYGLTETTSAILITPEGDDKPG 360  
QY 358 STGRPLPYIKAKVLDNATGKALGPGERGEICFQSEIMKGYNNPEATIDTIDKDWLHS 417  
Db 361 AVGVVLEHKGKMTKEIVDYVASQVTTAKKLKRGVGVFVDEVPKGLTGKLDARKIREILI 540  
QY 418 GOIYDDEGNFFIVDRKLKIKYGYQVAPAELENLLQHPSIADAGVTGVPDEFGGOL 477  
Db 421 GDIAYWDEHEFFIVDRKLSLIKYGQVAPAELESILLQHPNIFDAGVAGLPDDDAGEL 480  
QY 478 PAACVVLSESGKTLTEKEVQDFTAAQVTPTKHLRGGVWFVDSIPKGTGKLIRKELREIF- 536  
Db 481 PAAVVVLEHKGKMTKEIVDYVASQVTTAKKLKRGVGVFVDEVPKGLTGKLDARKIREILI 540  
QY 537 -AQRAPKSKL 545  
Db 541 KAKKGKSKL 550  
RESULT 5  
US-09-380-061B-21  
; Sequence 21, Application US/09380061B  
; Patent No. 6265177  
; GENERAL INFORMATION:  
; APPLICANT: SQUIRRELL, DAVID JAMES  
; WHITE, PETER JOHN  
; LOWE, CHRISTOPHER ROBIN  
; MURRAY, JAMES AUGUSTUS HENRY  
; TITLE OF INVENTION: ENZYME ASSAY FOR MUTANT FIREFLY LUCIFERASE  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NIXON & VANDERHYE P.C.  
; STREET: 1100 NORTH GLEBE ROAD  
; CITY: ARLINGTON  
; STATE: VIRGINIA  
; COUNTRY: U.S.A.  
; ZIP: 22201-4714  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/380,061B  
; FILING DATE: 25-Aug-1999  
; CLASSIFICATION: <unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/GB98/01026  
; FILING DATE: 7-APR-1998  
; APPLICATION NUMBER: GB 9707468.8  
; FILING DATE: 11-APR-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: SADOFF, B. J.  
; REGISTRATION NUMBER: 36,663  
; REFERENCE/DOCKET NUMBER: 124-725  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703)816-4000  
; TELEFAX: (703)816-4100  
; INFORMATION FOR SEQ ID NO: 21:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 550 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 21:  
US-09-380-061B-21

Query Match 56.3%; Score 1584.5; DB 4; Length 550;  
Best Local Similarity 55.5%; Pred. No. 1.4e-153;  
Matches 305; Conservative 103; Mismatches 137; Indels 5; Gaps 3;  
QY 1 MEE-ENIRHGERPRDIVHPGSAGQOOLYSLYKFPASPE--ATIDAHTNEVISYAQIFETS 57  
Db 1 MEDAKNIKKGPAPLEDEGTAGEQLHKAMKRYALVPGTIAFTDAHIEVNTIYAEEFMS 60  
QY 58 CRLAVSIEQYGLNENNVGVCSSENNINFPNPVLAALYIGIPVATSDMYTDOGELTGHINI 117  
Db 61 VRLAEMAKRYGLNTHRIWVCSSELSQPFMPVLFALFTGVAVAPANDIYNRELLNSMNI 120  
QY 118 SKPTIMFSSKALPILIRVQOONLSPIKVVVIDSMYDINGVECVSTFVARVYTDHTFDPLS 177  
Db 121 SPTVVVFSKGLQKILNVQKLPILQKIIIMDSKTDYQGFQSMYTFVTSHLPFGFNEYD 180  
QY 178 FTPKDFDPLEKIALIMSSGTTGLPKGVVLSHRSITIRFVHSRDIYGTTRVPTQSILSL 237  
Db 181 FVPSFDRDKTTIALTMNSSGSTGLPKGVALPHRTACVRFSHARDPIFGNQIIPDTAILS 240  
QY 238 VPFHAFGMFTTSLSYFVGLKVVMLKKFEGALFLKTIQNIKIPTIVAPPVVMFLAKSPL 297  
Db 241 VPFHGFGMFTTGLYLICGFRVLMYRFEELFLASLDQYKIQSALLVPTLFSFFAKSTL 300  
QY 298 VDQYDLSLSEVATGAGPLGKDVAAVAKRLKPLGLIOGYGLTETCCAVMTPHNAVKTG 357  
Db 301 IDKYDLSNLHEIASGAPLSKEVGAARFHLPGIRQGYGLTETTSAILITPEGDDKPG 360  
QY 358 STGRPLPYIKAKVLDNATGKALGPGERGEICFQSEIMKGYNNPEATIDTIDKDWLHS 417  
Db 361 AVGVVLEHKGKMTKEIVDYVASQVTTAKKLKRGVGVFVDEVPKGLTGKLDARKIREILI 540  
QY 418 GDIYDDEGNFFIVDRKLKIKYGYQVAPAELENLLQHPSIADAGVTGVPDEFGGOL 477  
Db 421 GDIAYWDEHEFFIVDRKLSLIKYGQVAPAELESILLQHPNIFDAGVAGLPDDDAGEL 480  
QY 478 PAACVVLSESGKTLTEKEVQDFTAAQVTPTKHLRGGVWFVDSIPKGTGKLIRKELREIF- 536  
Db 481 PAAVVVLEHKGKMTKEIVDYVASQVTTAKKLKRGVGVFVDEVPKGLTGKLDARKIREILI 540  
QY 537 -AQRAPKSKL 545  
Db 541 KAKKGKSKL 550  
RESULT 6  
US-08-122-520C-9  
; Sequence 9, Application US/08122520C  
; Patent No. 5639663  
; GENERAL INFORMATION:  
; APPLICANT: Crosby, William L.  
; APPLICANT: Datla, Raju S.S.  
; APPLICANT: Hammerlindl, Joseph K.  
; APPLICANT: Selvaraj, Gopalan  
; TITLE OF INVENTION: BIFUNCTIONAL GENETIC MARKERS  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: J. Wayne Anderson  
; STREET: Montreal Rd.  
; CITY: Ottawa  
; STATE: ON  
; COUNTRY: CANADA  
; ZIP: K1A 0R6  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/122,520C  
; FILING DATE: 11/29/92  
; CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:  
NAME: Anderson, J. Wayne  
REGISTRATION NUMBER: 28,158  
REFERENCE/DOCKET NUMBER: 10013-1A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (613) 993-3899  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 815 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-122-520C-9

Query Match 56.3%; Score 1584.5; DB 1; Length 815;  
Best Local Similarity 55.5%; Pred. No. 2.7e-153;  
Matches 305; Conservative 103; Mismatches 137; Indels 5; Gaps 3;  
QY 1 MEE-ENIRGERPRDIVHPGSAQQLYQSLYKFASFPE--AIIAHTNEVISYAOIFETS 57  
DB 1 MEDAKNIKKGPAPFPYLEDGTAGEQLHKAMKRYALVPGTIAFTDAHIEVNIYAEYFEMS 60  
QY 58 CRLAVSIEOYGLNENNVCVCSENNINFPVLAALYLGIPVATSDMYTDCGLTCHLNI 117  
DB 61 VRLAEAMKRYGLNTHRIVCVCSNSLQFMPVLGALFICGAVAPANDIYNERELNSMNI 120  
QY 118 SKPTIMFSSKKALPILRLVQOQLSFIKKVVIDSMYDINGVECVSTFVARYTDHTFDPLS 177  
DB 121 SPTVVFVSKKGLQILNVOKKLIPIQIIIMDSKTDYOGFOSMTFTVSHLPFGNEVD 180  
QY 178 FPKDFDPLEKIALIMSSSGTGLPKGVVLSHRSITIRFVHSRDIYGTTRTPQTSILS 237  
DB 181 FVPESFDRDKTIALIMNSSGTGLPKGVLPHTACVRFSHARDPIFGNQIIPDTAILS 240  
QY 238 VPFHAFGFTTSLYFVGLKVMMLKPFGLFKTIQNYKIPTIIVAPPVAVFLAKSPL 297  
DB 241 VPFHGFPGFTTGLYICGFRVVMYREELFLSLQDYKIQSALLVPTLFSFFAKSTL 300  
QY 298 VQOYDLSLSETEVATGAPIGKDVAAVAKRLKPLGIIOGYGLTETCCAVMITPHNAVKTG 357  
DB 301 IDKYDLSNLHETASGAPLSKEVGEAVAKRFLPGRIOGYGLTETTSAILITPEGDDKPG 360  
QY 358 STGRPLPYIKAKVLDNATKALGPCGERGICFQSEIMKGYNNPEATIDTIDKDGWLHS 417  
DB 361 AVGVVPPFEAKVVDLTGKTLGVNQRGELCVRGPIMSGVYNNPEATNALIDKDGWLHS 420  
QY 418 GDIGYDEGNFIVDRKLKELIKYQVAPAELENLLQHPISADAGVTGVPDDEFGGOL 477  
DB 421 GDIAWDEDEHFFIVDRKLSLKYQVAPAELESILLQHPNIFDAGVAGLPDDDAGEL 480  
QY 478 PAACVVLSESGKTLTEKEVODFTAAQVTPPKHLRGVGVVDSIPKGTGKLIKRELREIF- 536  
DB 481 PAADVLEHGKMTKEIYDVASQVTTAKKLRGGVVFVDEVPKGLTGKLDARKIREILI 540  
QY 537 -AORAPKSKL 545  
DB 541 KAKKGKSKL 550

## RESULT 7

US-08-718-425-2  
Sequence 2, Application US/08718425  
Patent No. 6132983  
GENERAL INFORMATION:  
APPLICANT: Love, Christopher R.  
APPLICANT: White, Peter J.  
APPLICANT: Murray, James A. H.  
APPLICANT: Squirell, David J.  
TITLE OF INVENTION: "LUCIFERASES"  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Nixon & Vanderhye

STREET: 1100 No. 6132983th Glebe Road, 8th Floor  
CITY: Arlington  
COUNTRY: U.S.A.  
ZIP: 22201-4714  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC Compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/718.425  
FILING DATE: 20-NOV-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Arthur R. Crawford  
REGISTRATION NUMBER: 25,327  
REFERENCE/DOCKET NUMBER: 124-539  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 816-4000  
TELEFAX: (703) 816-4100  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 550 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-718-425-2

Query Match 56.3%; Score 1583.5; DB 4; Length 550;  
Best Local Similarity 55.5%; Pred. No. 1.8e-153;  
Matches 305; Conservative 103; Mismatches 137; Indels 5; Gaps 3;  
QY 1 MEE-ENIRGERPRDIVHPGSAQQLYQSLYKFASFPE--AIIAHTNEVISYAOIFETS 57  
DB 1 MEDAKNIKKGPAPFPYLEDGTAGEQLHKAMKRYALVPGTIAFTDAHIEVNIYAEYFEMS 60  
QY 58 CRLAVSIEOYGLNENNVCVCSENNINFPVLAALYLGIPVATSDMYTDCGLTCHLNI 117  
DB 61 VRLAEAMKRYGLNTHRIVCVCSNSLQFMPVLGALFICGAVAPANDIYNERELNSMNI 120  
QY 118 SKPTIMFSSKKALPILRLVQOQLSFIKKVVIDSMYDINGVECVSTFVARYTDHTFDPLS 177  
DB 121 SPTVVFVSKKGLQILNVOKKLIPIQIIIMDSKTDYOGFOSMTFTVSHLPFGNEVD 180  
QY 178 FPKDFDPLEKIALIMSSSGTGLPKGVVLSHRSITIRFVHSRDIYGTTRTPQTSILS 237  
DB 181 FVPESFDRDKTIALIMNSSGTGLPKGVLPHTACVRFSHARDPIFGNQIIPDTAILS 240  
QY 238 VPFHAFGFTTSLYFVGLKVMMLKPFGLFKTIQNYKIPTIIVAPPVAVFLAKSPL 297  
DB 241 VPFHGFPGFTTGLYICGFRVVMYREELFLSLQDYKIQSALLVPTLFSFFAKSTL 300  
QY 298 VQOYDLSLSETEVATGAPIGKDVAAVAKRLKPLGIIOGYGLTETCCAVMITPHNAVKTG 357  
DB 301 IDKYDLSNLHETASGAPLSKEVGEAVAKRFLPGRIOGYGLTETTSAILITPEGDDKPG 360  
QY 358 STGRPLPYIKAKVLDNATKALGPCGERGICFQSEIMKGYNNPEATIDTIDKDGWLHS 417  
DB 361 AVGVVPPFEAKVVDLTGKTLGVNQRGELCVRGPIMSGVYNNPEATNALIDKDGWLHS 420  
QY 418 GDIGYDEGNFIVDRKLKELIKYQVAPAELENLLQHPISADAGVTGVPDDEFGGOL 477  
DB 421 GDIAWDEDEHFFIVDRKLSLKYQVAPAELESILLQHPNIFDAGVAGLPDDDAGEL 480  
QY 478 PAACVVLSESGKTLTEKEVODFTAAQVTPPKHLRGVGVVDSIPKGTGKLIKRELREIF- 536  
DB 481 PAADVLEHGKMTKEIYDVASQVTTAKKLRGGVVFVDEVPKGLTGKLDARKIREILI 540  
QY 537 -AORAPKSKL 545  
DB 541 KAKKGKSKL 550

Query Match	56.1%	Score 1576.5;	DB 4;	Length 550;
Best Local Similarity	55.3%	Pred. NO. 9.2e-153;		



Matches 304; Conservative 103; Mismatches 138; Indels 5; Gaps 3;

Qy 1 MEE-ENIRHGERPRDIVHPGSGAQQOYQSLYKFASFPE--AIIDAHTNEVSYAQIFETS 57  
Db 1 MEDAKNIKKGPAPFPYLEDGTAGEQLHKAMKRYALVPGTIAFTDAHIEVDITYAEYFEMS 60  
Qy 58 CRLAVSIOYGLNENNVGVCSENNINFFNPVLAALYLGIPVATNSDMYTDELTHLNI 117  
Db 61 VRLAEAMKRYGLNTHRVVCSNSLQFPVPGALFICGAVAPANDIYNRELLNSMI 120  
Qy 118 SKPTIMFSSKKALPLILRVQONLSFIKKVWIDSMYDINGVECVSTFVARVTDHFTDPLS 177  
Db 121 SQTIVVFSYKKGLOKILNVQKKLPDIIQIIIMDSKTDYQGFOSMTFTVTSHPGFEYD 180  
Qy 178 FTKPDFDPLEKIALIMSSGTTGPKGVLSHRSITIRFVHSRDPYIGTRTVPQTSILSL 237  
Db 181 FVPSEFDRDKTIALIMSSGTTGPKGVLPHTACVRFSHARDPIFGNQIIPDTAILSV 240  
Qy 238 VPFHAFGMFTTSLVYVGLKVMKKEGALFKTIQNYKIPTIVAPPVNVFLAKSPL 297  
Db 241 VPFHOGFGMFTTGLYICGFRVLMYREEEELFRLSLQDYKIQSALLVPTLFSFPAKSTL 300  
Qy 298 VDQYDLSLTVATGAPLGKDVAAEAVAKRLKPGIIOGYGLTETCCAVMITPHNAVKTG 357  
Db 301 IDKYDLSNLHIEASGAPLSKEVGEAVAKRFLPGLPIROGYGLTETTSAILITPEGDDKPG 360  
Qy 358 STGRPLPYIKAVLDNATGKALGPCERGEICFQSWIMKGYNNPEATIDTDKDGWLHS 417  
Db 361 AVGVVPEFEAKVVDLDTGKTLGVNORGLCVRGPMINSYNNPEATNALIDKDGWLHS 420  
Qy 418 GDIGYDEGNGFFIVDRKELIKYGYQVAPAELENLLQHPSIADAGVTGVPDEFQGL 477  
Db 421 GDIAWDEDEHFFIVDRKLSLIKYGQVAPAELESILLQHPNIFDAGVAGLPDDAGEL 480  
Qy 478 PAACVLESGLTKEVEQDFAAQVTPKHLRGVGVFVDSIPKGTCKLIRKELREIFA 536  
Db 481 PAVVVLEHGKMTKEIYDVASQVTTAKKLRGGVFEVDEVPKGLTKGLDARKIREILI 540  
Qy 537 -AORAPSKL 545  
Db 541 KAKKGGSKL 550

RESULT 10

US-08-354-240A-4  
; Sequence 4, Application US/08354240A  
; Patent No. 5670356  
; GENERAL INFORMATION:  
; APPLICANT: Sherf, Bruce A.  
; APPLICANT: Wood, Keith V.  
; TITLE OF INVENTION: MODIFIED LUCIFERASE  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dewitt Ross & Stevens, S.C.  
; STREET: 8000 Excelsior Drive, Suite 401  
; CITY: Madison  
; STATE: WI  
; COUNTRY: USA  
; ZIP: 53717-1914  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/354,240A  
; FILING DATE: 12-DEC-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sara, Charles S.  
; REGISTRATION NUMBER: 30,492  
; REFERENCE/DOCKET NUMBER: 34506.029  
; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 608-831-2100  
; TELEFAX: 608-831-2106  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 550 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-354-240A-4

Query Match 55.8%; Score 1568.5; DB 1; Length 550;  
Best Local Similarity 54.7%; Pred. No. 6.1e-152;  
Matches 300; Conservative 103; Mismatches 142; Indels 3; Gaps 2;

Qy 1 MEE-ENIRHGERPRDIVHPGSGAQQOYQSLYKFASFPE--AIIDAHTNEVSYAQIFETS 57  
Db 1 MEDAKNIKKGPAPFPYLEDGTAGEQLHKAMKRYALVPGTIAFTDAHIEVDITYAEYFEMS 60  
Qy 58 CRLAVSIOYGLNENNVGVCSENNINFFNPVLAALYLGIPVATNSDMYTDELTHLNI 117  
Db 61 VRLAEAMKRYGLNTHRVVCSNSLQFPVPGALFICGAVAPANDIYNRELLNSMI 120  
Qy 118 SKPTIMFSSKKALPLILRVQONLSFIKKVWIDSMYDINGVECVSTFVARVTDHFTDPLS 177  
Db 121 SQTIVVFSYKKGLOKILNVQKKLPDIIQIIIMDSKTDYQGFOSMTFTVTSHPGFEYD 180  
Qy 178 FTKPDFDPLEKIALIMSSGTTGPKGVLSHRSITIRFVHSRDPYIGTRTVPQTSILSL 237  
Db 181 FVPSEFDRDKTIALIMSSGTTGPKGVLPHTACVRFSHARDPIFGNQIIPDTAILSV 240  
Qy 238 VPFHAFGMFTTSLVYVGLKVMKKEGALFKTIQNYKIPTIVAPPVNVFLAKSPL 297  
Db 241 VPFHOGFGMFTTGLYICGFRVLMYREEEELFRLSLQDYKIQSALLVPTLFSFPAKSTL 300  
Qy 298 VDQYDLSLTVATGAPLGKDVAAEAVAKRLKPGIIOGYGLTETCCAVMITPHNAVKTG 357  
Db 301 IDKYDLSNLHIEASGAPLSKEVGEAVAKRFLPGLPIROGYGLTETTSAILITPEGDDKPG 360  
Qy 358 STGRPLPYIKAVLDNATGKALGPCERGEICFQSWIMKGYNNPEATIDTDKDGWLHS 417  
Db 361 AVGVVPEFEAKVVDLDTGKTLGVNORGLCVRGPMINSYNNPEATNALIDKDGWLHS 420  
Qy 418 GDIGYDEGNGFFIVDRKELIKYGYQVAPAELENLLQHPSIADAGVTGVPDEFQGL 477  
Db 421 GDIAWDEDEHFFIVDRKLSLIKYGQVAPAELESILLQHPNIFDAGVAGLPDDAGEL 480  
Qy 478 PAACVLESGLTKEVEQDFAAQVTPKHLRGVGVFVDSIPKGTCKLIRKELREIFA 537  
Db 481 PAVVVLEHGKMTKEIYDVASQVTTAKKLRGGVFEVDEVPKGLTKGLDARKIREILI 540  
Qy 538 ORAPSKL 545  
Db 541 KAKKGGKI 548

RESULT 11

US-08-354-240A-6  
; Sequence 6, Application US/08354240A  
; Patent No. 5670356  
; GENERAL INFORMATION:  
; APPLICANT: Sherf, Bruce A.  
; APPLICANT: Wood, Keith V.  
; TITLE OF INVENTION: MODIFIED LUCIFERASE  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dewitt Ross & Stevens, S.C.  
; STREET: 8000 Excelsior Drive, Suite 401  
; CITY: Madison  
; STATE: WI  
; COUNTRY: USA  
; ZIP: 53717-1914  
; COMPUTER READABLE FORM:



QY 537 -AORAPSKL 545  
Db 552 KAKRGKSKL 561

RESULT 13  
US-09-380-061B-20  
; Sequence 20, Application US/09380061B  
; Patent No. 6265177  
; GENERAL INFORMATION:  
; APPLICANT: SQUIRRELL, DAVID JAMES  
; WHITE, PETER JOHN  
; LOWE, CHRISTOPHER ROBIN  
; MURRAY, JAMES AUGUSTUS HENRY  
; TITLE OF INVENTION: ENZYME ASSAY FOR MUTANT FIREFLY LUCIFERASE  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NIXON & VANDERHVE P. C.  
; STREET: 1100 NORTH GLEBE ROAD  
; CITY: ARLINGTON  
; STATE: VIRGINIA  
; COUNTRY: U.S.A.  
; ZIP: 22201-4714  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/380.061B  
; FILING DATE: 25-Aug-1999  
; CLASSIFICATION: <unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/GB98/01026  
; FILING DATE: 7-APR-1998  
; APPLICATION NUMBER: GB 9707468.8  
; FILING DATE: 11-APR-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: SADOFF, B. J.  
; REGISTRATION NUMBER: 36,663  
; REFERENCE/DOCKET NUMBER: 124-725  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703)816-4000  
; TELEFAX: (703)816-4100  
; INFORMATION FOR SEQ ID NO: 20:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 547 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 20:  
US-09-380-061B-20

Query Match 54.68; Score 1534; DB 4; Length 547;  
Best Local Similarity 53.58; Pred. No. 2,1e-148;  
Matches 293; Conservative 106; Mismatches 145; Indels 4; Gaps 3;

QY 1 MEE-ENIRGERPRDIVHPGAGQOLYOSLYKFAFPE--AIIDAHTNEVSYAQIFETS 57  
Db 1 MEDAKNIMHGAPFPYLEDGTAGEQLHAKMKRYAQVPTIAFTDHAKEVNIYSYEFMA 60

QY 58 CRLAVSIEQYGLNENNVYGVCSENNINFPVLAALYIGIPVATSNMDYTGELTGLHNI 117  
Db 61 CRLAETMKRYGLQHHAIVCSSENSLOQFMPVCGALFVGCVASTNDIYNRELNSLSI 120

QY 118 SKPTIMFSKRALPILRVQONLSIKKVVVIDSMYDINGECVSTFVARTDHTFDPLS 177  
Db 121 SOPTIVSCSKRALQILGVOKLPITQIKIVILDSREDYMGKQSMYSFIESHLPAGFNEYD 180

QY 178 FTPKDPDPLEKIALWSSGGTGLPKGVLSHRSITIRFVHSRDPYIGTRVTPOTSILSL 237  
Db 181 YIPDSFDRETATALLWSSGGTGLPKGVLTQNVCRVFSHCRDPVFGNQIIPDTAILTV 240

QY 238 VPFHAEFGMETTILSYFVVLKVMMLKKEGALFLKTIQNYKIPTIVVAPPVMVFLAKSPL 297  
Db 241 IPFHGEGMETTILGYLTCGPRIVLMYRFEELFURSLQDYKIQSALLVPTLFSFAKSTL 300

QY 298 VDQYDLSLSEVATGGAPLKGDAEAVAKRLKPGIIGQYGLTETCCAVMITPHNAVKTG 357  
Db 301 VDQYDLSNLHEIASGGAPLAKEVGEAVAKRKLPGIRQYGLTETTSALIIITPEGDOKPG 360

QY 358 STGRPLPYIKAKVLONATGKALGPERGEICFQSEMIMKGYNNPEATIDTIDKDGWLHS 417  
Db 361 ACQKVPFESAIVDLDTGKTGLGVNQRGELCVKPGMIMKGYNNPEATISALIDKDGWLHS 420

QY 418 GDIGYDEDEGNFFIVDRKLKELIKYGVQVAPAELENLLQHPSTADAGVTPVDEFGQL 477  
Db 421 GDIAAYDKDGHFFIVDRKLSLIKYGQVPPAELESILLQHPFIPDAGVAGIPDPDAGEL 480

QY 478 PAACVWLESGKLTKEKEVQDFIAAQVTPTKHLRGVGVFVDSIPKGPCTGKLRKELREIFA 537  
Db 481 PAAVVLEEGKMTQEQVMDYVAGQVATSKRLRGVGFVDEVPKGLTGKIDGRKIREIL- 539

QY 538 ORAPSKL 545  
Db 540 MWGKSKL 547

RESULT 14  
US-09-111-752-5  
; Sequence 5, Application US/09111752  
; Patent No. 6074859  
; GENERAL INFORMATION:  
; APPLICANT: HIROKAWA, KOZO  
; APPLICANT: KAJIYAMA, NAOKI  
; APPLICANT: MURAKAMI, SEIJI  
; TITLE OF INVENTION: MUTANT-TYPE BIOLUMINESCENT PROTEIN, AND  
; TITLE OF INVENTION: PROCESS FOR PRODUCING MUTANT-TYPE LUMINESCENT PROTEIN  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
; ADDRESS: P.C.  
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400  
; CITY: ARLINGTON  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/111.752  
; FILING DATE: 08-JUL-1998  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OBLON, NORMAN F.  
; REGISTRATION NUMBER: 24,618  
; REFERENCE/DOCKET NUMBER: 7126-0009-0  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-413-3000  
; TELEFAX: 703-413-2220  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 552 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; ORIGINAL SOURCE:  
; ORGANISM: Luciola cruciata and Photinus pyralis  
US-09-111-752-5

```
Query Match 54.2%; Score 1524.5; DB 3; Length 552;
Best Local Similarity 54.6%; Pred. No. 2e-147;
Matches 298; Conservative 96; Mismatches 149; Indels 3; Gaps 2;

QY 3 ENIRHGERPRDIVPGSAGQOOLYOSLYKFASFPE-AIIDAHTNEVISYAIQIFETSCRLA 61
DB 7 DENIVVVGPKFPYPIEGSAGTQLRKYMERYAKLGAIAFTNAVTVGDVSYAEYLEKSCCLG 66
QY 62 VSIEOYGLNENNVGVCSENNINFPVLAALYLGIPIVATSNMVTDCGLTCHLNIKSPT 121
DB 67 KALQNYGLVDRGIALSCENCEEFFIPVIAGLFVGIVGAPTNEIYTLRELVHSLGISKPT 126
QY 122 INFSSKKALPLRLYQQNLSTFKVVVVDSDMYDINGECVSTFVARYTDHTDPLSFTPK 181
DB 127 IVFSSKKGLDKVITVQKTVTTIKTIVILDSKVDYRGYQCLDTFKRNTPPGQASSFKTV 186
QY 182 DDPLEKIALINSSSGTGLPGVVLSHRSLTIRVHSDPIYGRTPVQTSILSLVPPH 241
DB 187 EYDRKEQVALINSSSGTGLPGVQLTHENTVTRFESHARDPIFGNQIIPDTAILSVVPPH 246
QY 242 HAFGMFTTLYSVVGLKVVMKKFEGALFKTIQNYKIPTIVVAPPVVMVFLAKSPLVDQY 301
DB 247 HGFGMFTTLYGLICFRVVMVLTFKDEETFLKTLQDYKCTSVILVPTILFAILNKSELNNY 306
QY 302 DLSSLTEVATGAPLGKDVAAEAVAKRLKLPGLIQQYGLTETCCAVMITPHNAVKTGSTR 361
DB 307 DLSNLVEIASGAPLSKEVGEAVARRNLPVGRQCYGLTETTSAILIITPEGDDKPGASC 366
QY 362 PLPYTKAKVLDNATKALPGERGEICQSEMIMKGYNNPEATIDTIDKDGWLHSGDIG 421
DB 367 VYPLFKAKVIDLDTKSLGPNRRGEVCVKGPMLMKGYNNPEATKELIDEGWHLTGDI 426
QY 422 YYDEGNFFIVDRKLKELIKYQVAPAELENLLQHPISADAGVTGVDFEGGOLPAAC 481
DB 427 YDEKHFIVDRKSLIKYQVAPAELESILLQHPNIFDAGVAGLDDAGELPAAV 486
QY 482 VYLESGKLTKEVODFTAAQVTPTKHLRGGVVFVDSIPKPGTGLIRKELREIF--AQR 539
DB 487 VYLEHGKMTKEIYDVASQVTTAKKLRGGVVFVDEVPKGLTGKLDARKIREILIKAKK 546
QY 540 APKSKL 545
DB 547 GGSKL 552

RESULT 15
US-09-111-752-7
; Sequence 7, Application US/09111752
; Patent No. 6074859
; GENERAL INFORMATION:
; APPLICANT: HIROKAWA, KOZO
; APPLICANT: KAJIYAMA, NAOKI
; APPLICANT: MURAKAMI, SEIJI
; TITLE OF INVENTION: MUTANT-TYPE BIOLUMINESCENT PROTEIN, AND
; TITLE OF INVENTION: PROCESS FOR PRODUCING MUTANT-TYPE LUMINESCENT PROTEIN
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MATER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/111,752
; FILING DATE: 08-JUL-1998
; CLASSIFICATION: 435
```

```
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 7126-0009-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 552 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Luciola cruciata and Phontius pyralis
US-09-111-752-7

Query Match 54.1%; Score 1521.5; DB 3; Length 552;
Best Local Similarity 53.8%; Pred. No. 4e-147;
Matches 294; Conservative 101; Mismatches 148; Indels 3; Gaps 2;

QY 3 ENIRHGERPRDIVPGSAGQOOLYOSLYKFASFPE-AIIDAHTNEVISYAIQIFETSCRLA 61
DB 7 DENIVVVGPKFPYPIEGSAGTQLRKYMERYAKLGAIAFTNAVTVGDVSYAEYLEKSCCLG 66
QY 62 VSIEOYGLNENNVGVCSENNINFPVLAALYLGIPIVATSNMVTDCGLTCHLNIKSPT 121
DB 67 KALQNYGLVDRGIALSCENCEEFFIPVIAGLFVGIVGAPTNEIYTLRELVHSLGISKPT 126
QY 122 INFSSKKALPLRLYQQNLSTFKVVVVDSDMYDINGECVSTFVARYTDHTDPLSFTPK 181
DB 127 IVFSSKKGLDKVITVQKTVTTIKTIVILDSKVDYRGYQCLDTFKRNTPPGQASSFKTV 186
QY 182 DDPLEKIALINSSSGTGLPGVVLSHRSLTIRVHSDPIYGRTPVQTSILSLVPPH 241
DB 187 EYDRKEQVALINSSSGTGLPGVQLTHENTVTRFESHARDPIYGNQVSPGTAVLTVPVPH 246
QY 242 HAFGMFTTLYSVVGLKVVMKKFEGALFKTIQNYKIPTIVVAPPVVMVFLAKSPLVDQY 301
DB 247 HGFGMFTTLYGLICFRVVMVLTFKDEETFLKTLQDYKCTSVILVPTILFAILNKSELNNY 306
QY 302 DLSSLTEVATGAPLGKDVAAEAVAKRLKLPGLIQQYGLTETCCAVMITPHNAVKTGSTR 361
DB 307 DLSNLVEIASGAPLSKEVGEAVARRNLPVGRQCYGLTETTSAILIITPEGDDKPGASC 366
QY 362 PLPYTKAKVLDNATKALPGERGEICQSEMIMKGYNNPEATIDTIDKDGWLHSGDIG 421
DB 367 VYPLFKAKVIDLDTKSLGPNRRGEVCVKGPMLMKGYNNPEATKELIDEGWHLTGDI 426
QY 422 YYDEGNFFIVDRKLKELIKYQVAPAELENLLQHPISADAGVTGVDFEGGOLPAAC 481
DB 427 YDEKHFIVDRKSLIKYQVAPAELESILLQHPNIFDAGVAGLDDAGELPAAV 486
QY 482 VYLESGKLTKEVODFTAAQVTPTKHLRGGVVFVDSIPKPGTGLIRKELREIF--AQR 539
DB 487 VYLEHGKMTKEIYDVASQVTTAKKLRGGVVFVDEVPKGLTGKLDARKIREILIKAKK 546
QY 540 APKSKL 545
DB 547 GGSKL 552
```

Search completed: September 4, 2002, 15:03:56  
Job time: 205 sec